INDUSTRY AND IDENTITY

The Migration Linkage Between Economic and Cultural Change in 19th Century Britain *

Vasiliki Fouka†

Theo Serlin‡

January 2025

Abstract

Canonical modernization theories argue that industrialization fosters the adoption of common identities through labor migration. We study this process in England and Wales during the Second Industrial Revolution. Using microdata on individuals' names and migration decisions, we document uneven convergence towards a common identity. We develop and structurally estimate a model in which migration and cultural identities are inter-dependent, and which reproduces the observed patterns of identity change. Our estimates indicate strategic complementarities in identity choice, leading to coordination on focal identities. Which identities will be favored depends on patterns of industrialization across space. Where those identities will grow depends on costs of migration. Our model decomposes the contributions of different facets of modernization and elucidates how, where, and which national identities can emerge through bottom-up processes driven by economic change, in the absence of state coercion.

*The authors would like to thank Alberto Bisin, Rikhil Bhavnani, Volha Charnysh, Gary Cox, Elias Dinas, Amir Goldberg, Anna Grzymała-Busse, David Laitin, Biljana Meiske, Helen Milner, Jonathan Rodden, Kenneth Scheve, Alain Schläpfer, Marco Tabellini, Yanos Zylberberg and seminar participants at Jinan University, the University of Copenhagen Workshop on Migration and Culture, the Stanford GSB Workshop on Cultural Evolution and Organizational Theory, the 2024 APSA Annual Meeting and the 2024 ASSA Annual Meeting for helpful feedback.

†Stanford University, vfouka@stanford.edu

‡Princeton University, tserlin@princeton.edu

1 INTRODUCTION

In the 19th century, European states experienced both massive economic change and the replacement of local identities with homogeneous national identities. From Marx and Engels to modernization theory, a long line of scholars attributed identity change to industrialization, among other mechanisms through the channel of labor migration (Marx and Engels, 1978 [1848]; Deutsch, 1961; Gellner, 2006 [1983]). A flow of migrants of diverse cultural backgrounds from agricultural regions into industrial centers is thought to have promoted the breakup of local and particularistic cultural affiliations and their consolidation into overarching identities that led to the formation of modern nations. This paper asks, were these scholars right, and if so, how?

Theoretically, the implications of migration flows for the cultural landscape that follows industrial change are not straightforward. First, industrialization sets off opposing forces: it increases in-migration, which pushes toward cultural change, but also lowers out-migration, which promotes cultural retention. Second, while many accounts of nation-building emphasize the adoption of a single identity as a coordination device for migrants from diverse origins (Gellner, 2006 [1983], p.59; Weber, 1976, p.78), they offer little theoretical guidance on how this identity is chosen. Many accounts implicitly (Gellner, 2006 [1983]) or explicitly (Weber, 1976) view the state as homogenizing the populace from above and therefore favoring the culture of the center; yet this factor was absent from cases like 19th century Britain, where identity choice was driven to a larger extent by individual decisions. Finally, how the industrialization of one region influences migration and thus identity changes depends on what happens everywhere else. The geographic location of one industrial center relative to others affects where in-migrants will be drawn from and which identities they will coordinate on.

In this paper, we revisit this seminal literature by providing a formal microfoundation and empirical test of the implications of industrialization-induced migration for cultural change.

We build a model of migration and identity choice and estimate it using rich individual-level microdata on migration decisions and naming choices covering the universe of residents of England and Wales during the Second Industrial Revolution. In the model, agents choose where to migrate based on economic returns and their preference to be with others of the same culture. Cultural choices are made by agents' parents, who have tastes for specific cultures, but also wish to maximize their offspring's future expected welfare. This model allows us to study in a principled way the equilibrium effects of industrialization on local identities, accounting for interdependence across regions and offsetting changes in in- and out-migration in response to economic change. Additionally, this approach allows us to separate what are conceptually distinct components of modernization—industrial activity on the one hand, and changes such as connectivity and transport costs on the other—and examine their separate contributions to the identity changes we observe.

The equilibrium conditions of the model include regression-style equations that relate migration choices to the presence of coethnics and culture choices to implied migration opportunities. These give estimates of the key model parameters: the "homophily elasticity" which governs how migration choices respond to the presence of coethnics, and the "culture elasticity" which governs how cultural choices respond to migration opportunities. To account for the endogeneity of migration and cultural choices, we develop instrumental variables using the interaction between coal deposits and historic cultures; these instruments are uncorrelated with observed forms of cultural proximity. Given estimates of the key elasticities, we can back out unobserved taste parameters for specific destinations among migrants and specific cultures among parents. To validate the model, we show that these tastes correlate strongly with measures of cultural proximity, such as surname similarity and religious distance. We further show that the model rationalizes both the overall cultural change occurring in the period 1851-1911 and the observed heterogeneity in the patterns of change across space.

We provide several novel results. First, we estimate an important role for responsiveness of migration decisions to the presence of members of the same culture and of cultural choice

to future migration opportunities. Parents invest in identities in anticipation of children's location choices, which dampens diversity resulting from migration and pushes towards cultural homogeneity. Our estimates imply that identity choice displays strategic complementarities: homophily in location choice and parents responsiveness to children's expected welfare favor cultures with many members. This, in turn, means that economic development may promote the spread of one identity at the expense of others through coordination dynamics.

Second, the model provides insight into *which* identities are likely to increase in popularity and become focal as a result of industrialization. This prediction is based purely on individual-level decisions and strategic identity choices and abstracts from political mechanisms and centrally imposed homogenization through education, conscription or other channels commonly emphasized by the nation-building literature. In the period we examine, the identity of the Southeast of England spread because industrialization shifted to locations that were well-positioned to specifically attract more migrants from the Southeast relative to migrants from other regions. Those locations were in the South and East of the country and in the region around London. Had wage growth remained higher around the traditional centers of the First Industrial Revolution, other regional identities, such as the one of the North and the Midlands, could have been favored instead.

Third, the model opens the black box of modernization and disentangles the effects of its different components. Counterfactual estimations show that the decline in local cultures, particularly in the peripheries, is a result of falling migration costs and not industrial development per se. Had migration costs not dropped to the extent they did between 1851-1911 the population migrating outside their district of birth would have been 31% lower and the share of the population adopting the local culture would have been 16% higher. In contrast, industrial development does not unambiguously lead to cultural change. While falling migration costs generated a core-periphery gradient in cultural homogenization, peripheries with coal deposits were able to resist the tendency of other peripheries towards more rapid cultural loss. This is a result of the offsetting forces triggered by industrial growth on in-

and out-migration. Compared to the center of the country, in the periphery the lower rates of migration to other cultural clusters caused by the presence of coal were offset to a much smaller degree by in-migration from the outside. This helped preserve local cultures in regions like Wales and the Northeast of England.

More broadly, the response of cultural choice to local economic development depends on location. When industry develops in regions that can, due to their accessibility, pull in diverse migrants, adoption of a focal identity is more likely to ensue. Instead, when industrial development occurs in more distant locations, the retention of the local population fosters the preservation of local identities. While existing literature associates assimilation into national identities with industrial locations, and local identity preservation with agricultural regions lagging in terms of industrial adoption (Deutsch, 1966; Gellner, 2006 [1983]; Green, 2022), our results suggest that local identities can be preserved under industrialization, and do not only survive due to its absence.

2 RELATED LITERATURE

Our study relates to a broad literature on modernization and identity change. Classical accounts highlight many forces that drove the formation of unifying identities in Europe during the 19th and 20th centuries, such as education (Gellner, 1964), cultural diffusion through print media (Anderson, 1983), or the construction of infrastructure networks across markets (Weber, 1976), but have also emphasized the role of internal mobility (Deutsch, 1969; Weber, 1976; Gellner, 2006 [1983]). We theoretically microfound and empirically test this latter connection. In so doing, we confirm a role of industrialization-driven labor migration for the emergence of national identities, but also uncover new mechanisms.

Specifically, we show that industrialization and migration can lead to cultural change, but also sustain local identities, a finding not generally emphasized by extant theory. In this, we join recent literature that shows how modernization processes, such as railway construction, commonly associated with homogenization may also promote differentiating dynamics (Pengl

and Girardin, Forthcoming).[1] Our model shows that the spatial distribution of economic activity crucially determines whether local cultures will be lost or retained. Additionally, we contribute to this literature by showing that new national identities can emerge as a result of economic incentives and bottom-up rather than top-down processes of assimilation. As opposed to other well-known cases of identity transformation such as France (Weber, 1976; Blanc and Kubo, 2024) Britain did not engage in state-led nation-building efforts. There was no conscription, and education was administered at the local level. Our findings relate to those of recent studies showing that individual-level decisions and interactions can account for cultural homogenization and nation-building even in the absence of top-down coercion (Green, 2019, 2022; Bazzi et al., 2019).

More broadly, we contribute to the literature on the determinants of identity choice. Much work across disciplines considers identity to be an individual decision responding to economic and non-economic incentives (Waters, 1990; Laitin, 1995; Lazear, 1999; Shayo, 2009; Sambanis and Shayo, 2013). We develop a model in which economic incentives influence cultural choices by affecting where people migrate, and whether they seek to coordinate on culture with the people already there. The model combines elements of the framework of Bisin and Verdier (2001), in which parents choose their offspring's culture based on expected returns and their own preferences, with insights from Laitin (1994), who views the choice of culture as a coordination game (see also Laitin (2007)). Our quantitative estimates suggest an important role for both coordination dynamics and strategic parental decisions in choosing identities most likely to benefit their offspring.

Lastly, we apply empirical methods from a growing body of work on quantitative spatial models to the study of cultural change. Recent work in urban and international economics

[1]Studies have pointed out that uneven industrial development can drive differentiation and strengthen pre-existing cultural cleavages through mechanisms such as economic domination (Hechter, 1977) or status considerations and elite incentives (Gellner, 1964). Second-generation modernization literature that focused on Africa also proposed that modernization may strengthen, instead of eradicating, narrow group identities such as ethnicity (Bates, 1974), with mixed empirical evidence (Robinson, 2014; Pengl, Roessler and Rueda, 2022). This strand of research examines yet another set of mechanisms distinct from those we focus on and more related to group competition over the gains of modernization.

has developed tractable models of economic geography (Allen and Arkolakis, 2014; Redding and Sturm, 2008). These models can rationalize observed migration and commuting flows, and, when calibrated to observed data, generate accurate predictions of how changes in trade costs alter the spatial distribution of economic activity (Redding and Sturm, 2008; Ahlfeldt et al., 2015). We use this framework to study how changes in economic activity and migration costs affect the map of local identities. In addition, Monte, Redding and Rossi-Hansberg (2018) and Caliendo et al. (2018) show that the spatial structure of the economy generates heterogeneity in the impact of shocks to labor demand and productivity on employment and output. We show that the same is true of the relationship between economic growth and cultural choices.

3 HISTORICAL CONTEXT AND DATA

3.1 *England and Wales during the Second Industrial Revolution*

Our context of study is England and Wales in the period 1851–1911. During this time, Britain underwent rapid economic transformation, marked by a dramatic change in the sectors and spatial patterns of economic activity. Technological developments such as the Bessemer Process for making steel (1856), the Solvay Process for making ammonia (1861), and the replacement of wooden sailing ships with steel steamships in the 1870s (Crouzet, 1982), led to a growth of the steel, chemicals, and engineering and secondary metals industries which came to rival the textiles industry. Wool manufacturing made up 9% of manufacturing employment in 1851, but 4% in 1911; machinery manufacturing increased from 3% to 10% of manufacturing employment.[2]

The growth of new industries altered the spatial distribution of economic activity. Figure 1 shows the share of employment in manufacturing across regions in 1851, and the change in the log number of manufacturing workers between 1851 and 1911. In 1851, manufacturing was

[2]These figures were calculated from I-CeM census data (Schürer and Higgs, 2014). Figure A.1 shows employment in different manufacturing industries in 1851 and 1911.

concentrated in Lancashire, the center of the cotton industry. Between 1851 and 1911, rural areas in the Southwest, East Anglia, Wales, and the North declined and already industrialized parts of Lancashire and the Midlands experienced continued growth. Growth did not just conform to existing patterns of development. The new steel and metals industries were located near major coalfields in South Wales and the Northeast. During this period London and the Southeast also experienced rapid growth, in part due to new industries locating close to investors in the City of London (see also Geary and Stark 2015).

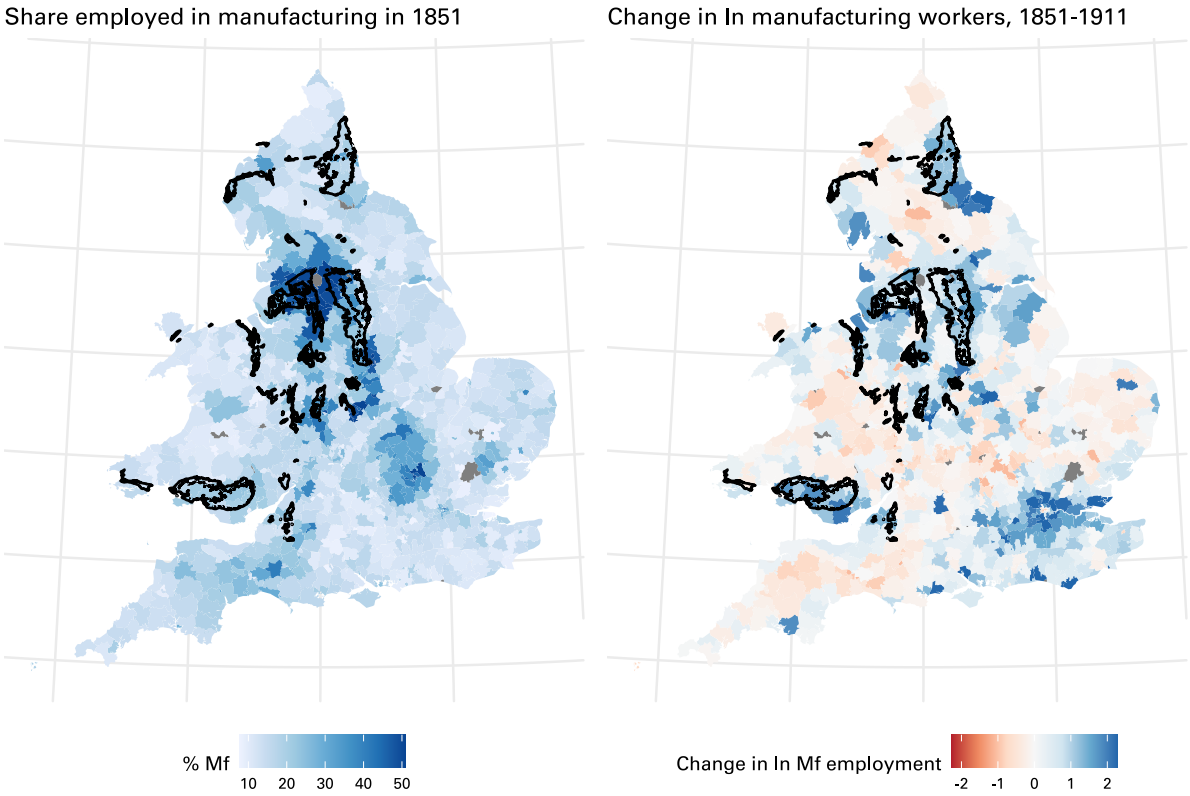


Figure 1: Industrialization in 1851, and growth over the 1851–1911 period

Black lines outline coal deposits.

3.2 *Data*

Our main data source is the I-CeM full count microdata from the Census of England and Wales, 1851–1911 (Schürer and Higgs, 2014). Two features of this data make it especially

useful for our purposes. First, the census recorded granular information on birthplaces, which allow us to trace out lifetime migration. Second, we have access to full names in the 1911 census, which we use to measure cultural choices, and the surnames of household heads in earlier censuses, which we use to cluster locations into cultural groupings.

MIGRATION We trace lifetime migration using information on individuals' birthplaces and locations of residence. Our unit of analysis throughout is the registration district, fixed to 1851 boundaries. Registration districts were administrative units created in the 1830s to register births, deaths and marriages.[3] The advantage of focusing on registration districts rather than smaller units such as parishes is that other social and economic data from this period which we use to validate our model, for instance attendance at churches of different religious denominations, were reported at the district level. We link birthplaces to registration districts using crosswalk files from Day (2016) and a GIS from Satchell et al. (2016).

During the period we examine, the volume of migration flows increased drastically. Figure 2 shows the shares of population living at different distances from their birthplaces in successive censuses. Between 1851 and 1911, the share living outside their district of birth increased by 31%, and people tended to migrate over longer distances.

LOCAL CULTURES Historically, Britain has been characterized by a high degree of diversity in local cultures and customs, driven not only by differences between its constitutive nations, but also differences within England (see for instance Homans (1969)). Variation in local geography and soil types, as well as centuries of limited communication across regions had generated "a patchwork in which uncertain areas of Welshness, Scotishness and Englishness were cut across by strong regional attachments, and scored over again by loyalties to village,

[3]In a few cases a registration district falls into more than one county (the largest administrative unit in this period). To facilitate the use of county-level data, we split registration districts falling into multiple counties, so that each unit falls into one county. We do so by allocating parishes to registration district-county units.

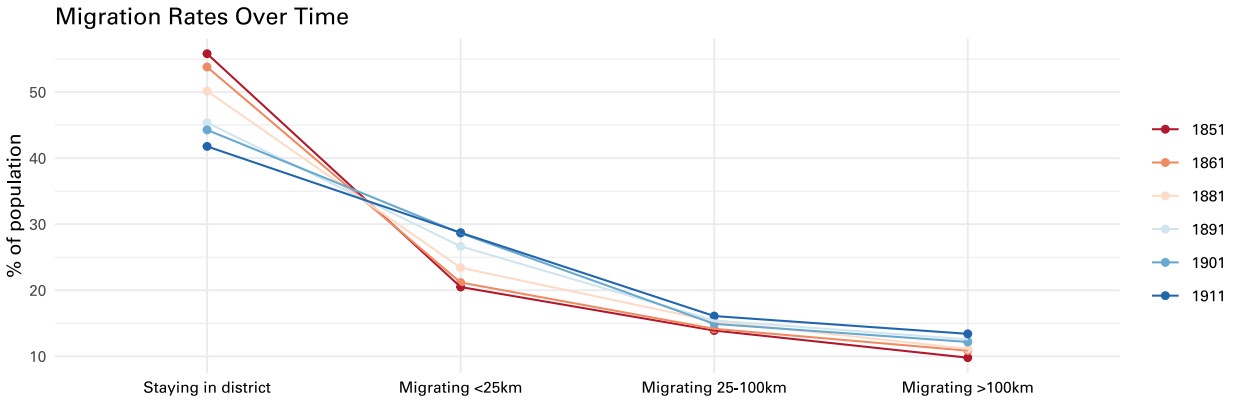


Figure 2: Increasing rates of migration 1851–1911

The figure plots the share of people living at different distances from their district of birth at the time of the census, in the 1851–1911 censuses.

town, family and landscape" (Colley, 2005, p.17). To quantify the distribution of these regional identities, we use a data-driven way to allocate districts to cultural "clusters" or regions that display high cultural similarity.

Our approach uses information on the surnames of household heads born before 1800, recorded in the 1851 census. We record the share born in each district with each surname. This data predates the Second Industrial Revolution.[4] The logic of examining surnames is that they trace out historic patterns of migration. For instance Porteous (1982) examines the surname "Mell," which likely derives from the Danish word for meal. He finds that it has been concentrated since at least the 16th century in the North East of England, which experienced large scale Norse settlement in the early middle ages. Kandt, Cheshire and Longley (2016) find that geographical clusters of surnames in the UK correspond to geographical clusters estimated using genetic data.

Given a matrix in which each row is a vector of surname shares for a given district, we use the spectral clustering algorithm developed by John et al. (2020) to cluster districts.

[4]In 1800, the First Industrial Revolution was underway, but growth was slow—Antràs and Voth (2003) estimate productivity growth of 0.2% per annum 1770–1800—and confined to specific industries like textiles (Mokyr, 2008). 1800 predates the large scale adoption of steam engines and the invention of railways. The industries that accounted for growth over the second half of the 19th century—especially steel, secondary metals industries like steel shipbuilding, and chemicals—did not yet meaningfully exist.

This algorithm first calculates a kernel similarity matrix between districts and then runs a Gaussian Mixture Model on the eigenvectors of the kernel matrix. It also uses these eigenvectors to estimate the optimal number of clusters. The resulting clusters, depicted in Figure 3, correspond to historical regions. For instance the W region of Figure 3 corresponds to the distinctive "West Country," and combined with the S E region traces out the boundaries of the Anglo Saxon Kingdom of Wessex, while the E region corresponds to East Anglia.

These regional boundaries were also predictive of behavior. In addition to being physically closer, two locations allocated to the same cluster on average had more similar patterns of religious attendance, greater overlap in the linguistic features of medieval texts originating from the locations, greater migration flows and a larger number of marriages between individuals born in the two locations (Table B.1). The number of migrants between two locations was 35% higher if the two locations belonged to the same cluster, adjusting for the role of distance and religious similarity; the number of marriages between people born in two locations was 11% higher even after adjusting for distance, religious similarity, and migration flows.[5] Clearly, there exist alternative ways to divide up England and Wales according to culture. In Appendix J.6, we verify that our results are robust to alternative cultural clusters based on linguistic profiles and administrative geographies.[6]

CULTURAL CHOICE We proxy for cultural choices using first names. We take names to be a summary indicator of other changes in markers and behaviors that could be reflective of identity change, but are harder to measure in our context. Existing scholarship indicates that names are a good proxy for latent attributes such as values and group attachments, and strongly correlate with behavioral measures of identity (Fouka, 2020; Bazzi, Fiszbein and

[5]Day (2023) uncovers a similar set of cultural regions by applying a community detection algorithm on migration flows between 1851 and 1911. Migration is a choice outcome in our model, which is why we prefer to rely on pre-determined cultural markers like surnames to identify cultural clusters prior to 1851.

[6]We consider both of these alternative approaches to be inferior to reliance on surnames. Linguistic profiles come from written records and do not capture the granularity of spoken local dialects, while boundaries of administrative units are not necessarily drawn with cultural differences in mind.

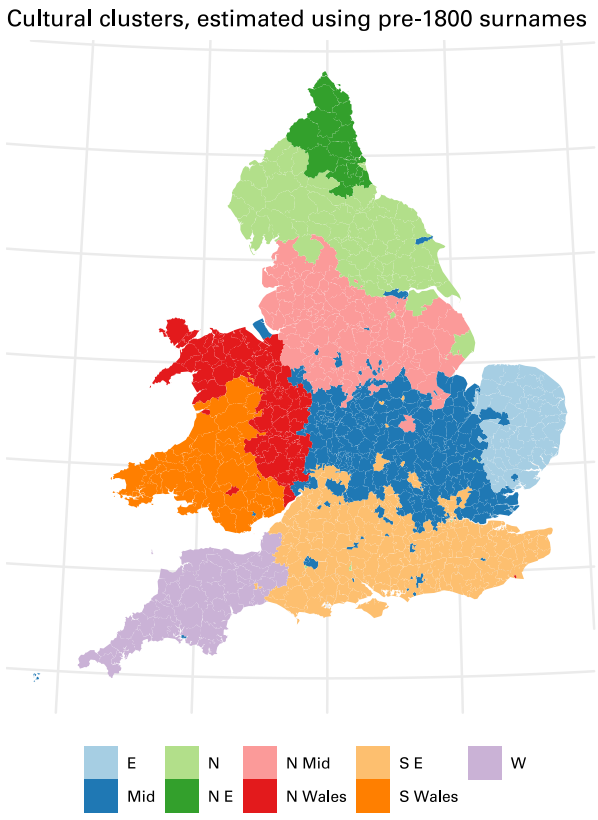


Figure 3: Map of cultural clusters based on surname frequencies

Clusters from running spectral clustering algorithm on surname frequencies of individuals present in the 1851 census who were born before 1801.

Gebresilasse, 2020).

We allocate names to cultural groups using data on the earlier generation in our data, those born between 1841 and 1860, who were aged 51–70 in 1911. We can then examine the types of names—in terms of their local identity content—that were given to the younger generation.

Using the frequencies of names among the 1841–1860 cohort in different registration districts, and an allocation of registration districts to a set of cultural clusters, we calculate culture name scores for each name i and each cluster k:

$$\textit{Culture Name Score}_{i,k} = \frac{P(\text{name} = i|\text{culture} = k)}{P(\text{name} = i|\text{culture} = k) + P(\text{name} = i|\text{culture} \neq k)}$$

where culture in the above expression refers to whether an individual was born in a given cultural cluster.[7] Name scores of this variety have been used widely to study racial identity (Fryer Jr and Levitt, 2004), immigrant assimilation (Abramitzky, Boustan and Eriksson, 2014; Fouka, 2019), and nation-building (Bazzi et al., 2019). Table B.2 reports the highest-scoring names for each cluster. These names trace out linguistic divides—Welsh names like Gwenllian and Llewellyn score highest in the Welsh clusters. High-scoring names also reflect local histories. Cuthbert, the namesake of the patron saint of Northumbria in the northeast of England, is the highest scoring name for the Northeast English cluster. Sutcliffe, a surname which originates from a number of villages in Yorkshire, is the highest scoring name for the North-Mid English cluster.

With culture name scores at hand, we can quantify identity change by examining names given to subsequent generations and the cultural clusters those names were associated with. This allows us to trace the popularity of different local cultures over time. For our main analyses, we allocate individuals born 1861–1895, who were adults at the time of the census, to the cluster for which their name receives the highest score. For instance, we consider the name David—the patron saint of Wales—to reflect the South Wales culture, because its name score for that cluster is 0.92, while its two next highest name scores are 0.67 for the North Wales cluster and 0.44 for the Eastern England cluster. Combined with information on birthplaces and places of residence, this approach gives origin-by-destination-by-culture migration flows.

Table B.3 shows that individuals given names associated with a culture were more likely to migrate to that region, more likely to marry people also given names associated with that culture, and more likely to give their children names associated with that culture. These associations hold even adjusting for the propensity of individuals in a given location to migrate, marry in, or assign their children to a particular culture. Table B.4 shows that those with higher name scores for the two Welsh clusters were more likely to speak Welsh and less

[7]We standardize name spellings using Metaphone codes.

likely to speak English. Name choices by parents can then reasonably be inferred to correlate with other meaningful acculturation decisions, such as language investments, and therefore represent an informative proxy of overall identity change.

We document three key facts about the process of cultural change during this period. First, there was significant cultural homogenization. Figure 4 maps the cluster with the highest average name score among those born in each registration district in the periods 1851–1860 and 1901–1910. The culture of the Southeastern cluster increased in popularity relative to other clusters; entire cultures disappeared from the map by the 1900s. Cultural homogenization, as measured with naming, in the late 19th century coincides with other forms of observed cultural homogenization. An overview of linguistic change in 19th century England notes that regional dialects declined in prevalence from the 1870s onwards, and were replaced by a linguistic standard based on the dialect of the area around London (Görlach, 1999). The apparent decline of dialects motivated intellectuals, like the novelist Thomas Hardy, to begin cataloging regional dialects. They linked the death of dialect to migration out of rural regions.

Second, the spread of the Southeast culture and the loss of home cultures was more pronounced in the peripheries. Table 1 column (1) regresses the change in the prevalence of the Southeast culture against distance to London. A 100km increase in distance to London corresponds to a 20% larger increase in the prevalence of the Southeast culture. Column (2) shows a corresponding decline in the prevalence of local cultures further from London.

Third, local industrialization had a heterogeneous effect on cultural change. Column (3) adds an indicator for whether the district contains coal deposits, which contributed to industrialization. On average, districts with coal experienced larger declines in local cultures. Peripherality moderated this relationship. Column (4) interacts coal and distance. This interaction term is positive and slightly larger in magnitude than the coefficient on distance. The presence of coal is associated with a decline in local cultures in districts near the center of the country; in the periphery instead, coal deposits sustained relative cultural holdouts.

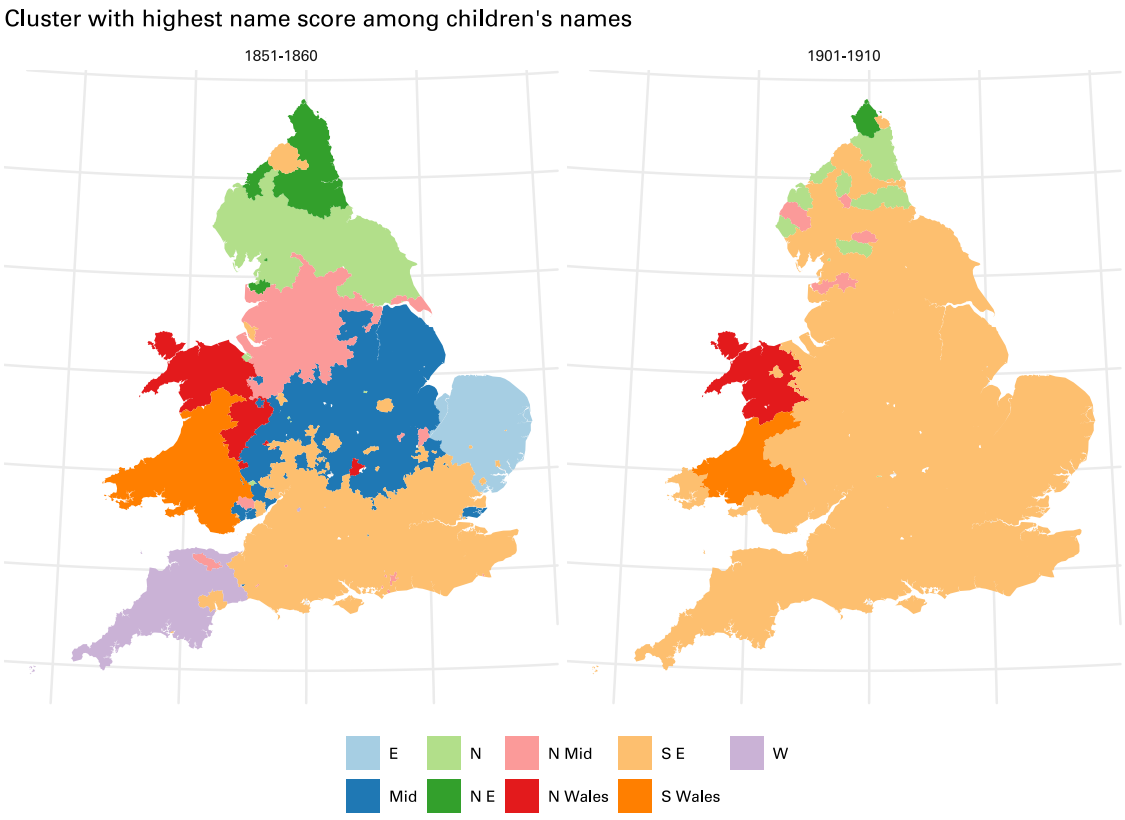


Figure 4: Changing spatial distribution of culture-specific names

These maps show the cluster with the highest name score among children born in each district, 1851–1860 (left) and 1901–1910 (right). We calculate the average name score for each cluster for those born in the district, and then record the cluster with the highest average score.

4 A MODEL OF MIGRATION AND IDENTITY CHOICE

We consider simultaneous migration and identity decisions in a spatial equilibrium model. In the model, agents, who are distinguished by their location of origin and culture, decide where to live. They choose locations in part due to economic incentives, and in part due to a preference to reside with others of the same culture. While we anticipate, and indeed find, that this preference is positive, our model is relatively general and treats the direction and magnitude of that preference as a quantity to be estimated. Agents also have culture- and individual-level idiosyncratic preferences for specific locations. Their parents choose to raise them in a particular culture, not knowing fully what these individual-level preferences will be. We model the decision of parents as a choice over the assignment of an identity

	Δ ln S E	Δ ln home culture		
	(1)	(2)	(3)	(4)
Distance to London (100km)	0.186	−0.168	−0.137	−0.169
	(0.008)	(0.016)	(0.018)	(0.018)
Contains coal			−0.141	−0.607
			(0.047)	(0.140)
Distance to London × coal				0.194
				(0.057)
Intercept	0.270	0.291	0.286	0.329
	(0.010)	(0.026)	(0.026)	(0.026)
N	827	798	798	798
R^2	0.545	0.195	0.213	0.239

This table presents evidence of the relationship between distance to London, coal, and the strength of the Southeast English and home cultures, using data on naming for the 1851–1860 and 1901–1910 cohorts. Regressions are at the district-level. The dependent variable in (1) is the change in the log share allocated names most associated with the Southeast English cluster, between those born 1851–1860 and 1901–1910, in (2)–(4), the change in the log share allocated names most associated with the home culture. Coal presence is an indicator that the district contains coal deposits, distance to London is distance from the district centroid to the City of London, in hundreds of kilometers. Models are weighted by the number allocated names in the 1851–1860 generation. Robust standard errors in parentheses.

Table 1: Relationship between coal, distance to London, and changing popularity of cultures

marker—corresponding to name choice in our data—which, however, we take to proxy for additional cultural investments the parents may make in a specific identity, such as in language or education. Parents care about the expected welfare of their child under different cultures—though we infer from the data the sign and magnitude of this welfare weight—but also have their own preferences for specific cultures.

4.1 *Setup*

There are N regions, and K cultures. We will use subscripts d (for destination), o (for origin), and j to index regions, and superscripts k and l to index cultures. There are two sets of agents: parents and children. Parents choose which culture to raise their offspring in; the younger generation then, having received an assigned identity from their parents, decides where to live and work. When choosing where to migrate, the utility person i of culture k

born in location o receives from migrating to destination d is

$$u_{odi}^k \equiv v_d \delta_{od} \left(m_d^k\right)^{\frac{\alpha}{\theta}} \xi_d^k \varepsilon_{di}.$$

v_d is the real wage in d, an attribute of the location that is experienced equally by members of all cultures.[8] δ_{od} is an inverse cost of migrating from o to d, and captures the idea that it is easier to migrate from a given origin to some destinations than others. m_d^k is the number of people who end up in d drawn from culture k, and α is the "homophily elasticity," the elasticity of migration choices to the population of the same culture. If α is positive, people have a preference for locating in places with more people of the same culture as them. ξ_d^k is a taste parameter, known to the agents though unobserved in the data, for location d for all members of culture k. This "cultural amenity" variable represents unmodeled preferences for certain places by members of certain cultures, distinct from the benefit of being with other members of the same culture. One such preference would be a desire to live in an ancestral homeland. ε_{di} is a preference shock drawn iid from a Fréchet distribution with shape θ over individual-by-destination pairs. This preference shock accounts for individual heterogeneity in migration decisions.[9]

Each agent migrates to the place that gives her the greatest utility, inclusive of her preference shock. It follows from the properties of the Fréchet distribution (McFadden, 1974; Hsieh et al., 2019; Bryan and Morten, 2019), that the number of members of culture k in o migrating to d is given by

$$m_{od}^k = \frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\sum_{j=1}^N \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha} n_o^k, \tag{1}$$

[8]While we do not explicitly model the economy, one can think of v_d as capturing factors like productivity and resource endowments that influence output per worker. This functional form could emerge if in each location a homogeneous and freely-traded numeraire good is produced using labor under perfect competition and constant returns to scale, with productivity v_d.

[9]The θ parameter governs the responsiveness of migration decisions to the varying utilities of different destinations. Our empirical strategy does not separately identify θ, but we do not need to do so to estimate counterfactuals.

where n_o^k is the number of people originating in o with culture k. This expression implies that more people will migrate to a location if wages are high (v_d is large), if it is near to places with large numbers of potential migrants ($\delta_{od}^{\theta} n_o^k$ is large), if it is more popular specifically with the cultures of these migrants ($\left(\xi_d^k\right)^{\theta}\left(m_d^k\right)^{\alpha}$ is large), and if there are few other compelling locations competing for the same migrants (the denominator is small).

It also follows from the properties of the Fréchet distribution that expected utility—prior to the realization of the ε_{di} shock—for members of culture k in o, is

$$\left(\sum_{j=1}^{N}\left(v_j\delta_{oj}\xi_j^k\right)^{\theta}\left(m_j^k\right)^{\alpha}\right)^{\frac{1}{\theta}}\Gamma\left(\frac{\theta-1}{\theta}\right),$$

where $\Gamma(\cdot)$ is the Gamma function. Ignoring the Γ constant, we define the variable Ω_o^k:

$$\Omega_o^k \equiv \sum_{j=1}^{N}\left(v_j\delta_{oj}\xi_j^k\right)^{\theta}\left(m_j^k\right)^{\alpha}. \tag{2}$$

An agent's parents choose to assign her to one of K cultures not knowing her vector of ε_{di} shocks. They make this cultural decision trading off the benefits of migration and cultural homogeneity against a place-specific attachment to different cultures. This assumption is similar to, but simpler than, the "imperfect empathy" of Bisin and Verdier (2001). Instead of parents evaluating their children's choices with their own utility function, they simply have preferences over their children's cultural identities that depend partly on the children's welfare and partly on the parent's tastes. More formally, the utility the parent of agent i in o receives from choosing culture k is

$$\bar{u}_{oi}^k \equiv \left(\Omega_o^k\right)^{\frac{1}{\theta}}\left(\psi_o^k\right)^{\frac{1}{\varphi\theta}}\iota_i^k$$

where ψ_o^k is the "cultural transmission taste," a place-specific bias that parents have for assigning their children a given culture, and ι_i^k is an individual-by-culture Fréchet shock, with shape $\varphi\theta$. It follows again from the properties of the Fréchet distribution that the share

choosing culture k in o is

$$\sigma_o^k = \frac{\left(\Omega_o^k\right)^{\varphi} \psi_o^k}{\sum_{l=1}^{K} \left(\Omega_o^l\right)^{\varphi} \psi_o^l}. \tag{3}$$

Here φ is the "culture elasticity," the parameter that determines how responsive cultural choices are to migration opportunities. A larger value of φ implies that the distribution of ι_i^k is less dispersed, meaning that differences in individual ι_i^k shocks for different cultures will be small in relation to Ω_o^k, and consequently choices will largely depend upon Ω_o^k, the expected utility experienced by the child raised in that culture. As φ tends towards infinity, we would expect all parents in a given location to choose the same culture for their children.

4.2 *Equilibrium*

An equilibrium in our model is a vector of culture-by-origin choices and culture-by-origin-by-destination migration flows such that equations (1) and (3) are satisfied. This implies the assumption of optimal behavior for each parent, fixing the actions of all other parents.

5 TAKING THE MODEL TO THE DATA

In the model, these migration flows and cultural choices are generated by agents optimizing according to preferences defined by the structural parameters $\alpha, \varphi, v_d^{\theta}, \delta_{od}^{\theta}, \left(\xi_d^k\right)^{\theta}$, and ψ_o^k, for $o, d \in \{1, ..., N\}$, and $k \in \{1, ..., K\}$. We treat the observed migration flows and cultural choices as an equilibrium realization of the model, and back out the structural parameters in three stages. Table 2 provides an overview of the different parameters and the estimation steps we follow to recover them from our data.

Table 2: Structural Parameters

	Reference	Interpretation	Level	Estimation
α	Homophily elasticity	Responsiveness of migration decisions to other members of the same culture	Universal	Two-stage least squares
φ	Culture elasticity	Responsiveness of cultural choices to migration opportunities	Universal	Instrumental variables Poisson regression
δ_{od}^{θ}	Inverse migration cost	Inverse cost of migrating from o to d	Origin-by-destination	Gravity regression of migration flows against distance and fixed effects
v_d^{θ}	Real wage	Non-cultural incentive to migrate to location d	Destination	Backed out from fixed effects in gravity regression
$\left(\xi_d^k\right)^{\theta}$	Cultural amenity	Culture k-specific taste for migrating to location d	Culture-by-destination	Backed out from fixed effects in gravity regression
ψ_o^k	Cultural transmission taste	Taste of parents in o for assigning children culture k	Culture-by-origin	Backed out from cultural choices in origin locations
Ω_o^k	Expected utility	Expected value of belonging to culture k if born in location o	Culture-by-origin	Function of α, v_d, δ_{od}, ξ_d^k and migration flows

5.1 *Estimating Migration Pull Factors and Costs*

In the first stage, we estimate the bundle of features that make a given location attract migrants from a given culture. Taking logarithms of (1) gives

$$\underbrace{\ln m_{od}^k}_{\ln \text{ migrants from } o \text{ to } d \text{ of culture } k} = \underbrace{\ln\left(v_d^\theta \left(\xi_d^k\right)^\theta \left(m_d^k\right)^\alpha\right)}_{\text{Destination-culture FE}} + \underbrace{\theta \ln \delta_{od}}_{\ln \text{ distance}} + \underbrace{\ln\left(n_o^k/\Omega_o^k\right)}_{\text{Origin-culture FE}}.$$

This equation implies that in a regression of log migration flows against some measure of migration costs and origin-by-culture and destination-by-culture fixed effects, the destination-by-culture fixed effect recovers the (log) expected utility of moving to a given location for members of each culture. We parameterize migration costs as a log-linear function of geographic distance and an indicator that the origin equals the destination, $\delta_{od}^\theta = \text{distance}_{od}^{\beta_1} \exp\left(\beta_2 \mathbf{1}_{\{o=d\}}\right)$. This choice follows scholarship on the gravity structure of migration, which indicates that migration flows are strongly related to distance. Inserting this parameterization of δ_{od}^θ into the equation above gives an estimating equation:

$$\ln m_{od}^k = \gamma_d^k + \beta_1 \ln \text{distance}_{od} + \beta_2 \mathbf{1}_{\{o=d\}} + \gamma_o^k + \varepsilon_{od}^k.$$

We estimate this model by Poisson Pseudo-Maximum Likelihood (Silva and Tenreyro, 2006; Fally, 2015). The error term ε_{od}^k here represents unobserved shocks to origin-by-destination migration costs. We assume these are transitory and do not affect choices of culture. The destination-by-culture fixed effect γ_d^k estimates $\ln\left(v_d^\theta \left(m_d^k\right)^\alpha \left(\xi_d^k\right)^\theta\right)$.

5.2 *Disaggregating Migration Pull Factors*

This estimated bundle of features that attract migrants to a location includes both exogenous cultural amenities ξ_d^k, and endogenous factors related to the presence of other members of a given culture in that location, $\left(m_d^k\right)^\alpha$, as well as the real wage that attracts members of all cultures to a location, v_d^θ. In the second stage we disentangle these three components. Writing

the definition of γ_d^k in separate terms gives a regression equation:

$$\underbrace{\gamma_d^k}_{\text{Destination-culture FE}} = \alpha \underbrace{\ln m_d^k}_{\ln \text{ destination-culture population}} + \underbrace{\ln v_d^\theta}_{\text{Destination FE}} + \underbrace{\ln \left(\xi_d^k\right)^\theta}_{\text{Error}}. \tag{4}$$

In a regression of the estimated destination-by-culture fixed effects against the log post-migration population of a given culture in a given destination and a destination fixed effect, the coefficient on log migration flows recovers α, the homophily elasticity, the destination fixed effect recovers (log) v_d^θ, the real wage, and the regression residual recovers the log of $\left(\xi_d^k\right)^\theta$, the cultural amenity.

Estimating (4) is complicated by migration flows being endogenous to cultural amenities. If a feature of place d makes it more attractive to members of a culture k (increasing ξ_d^k), more members of k will migrate to d, increasing m_d^k. Thus the independent variable $\ln m_d^k$ will be positively correlated with the error term $\ln \left(\xi_d^k\right)^\theta$. For instance, because members of the Southeast culture had a taste for migrating to places with Church of England churches (a form of cultural amenity), the presence of members of the Southeast culture (m_d^k) is correlated with historical Church of England attendance (ξ_d^k) (Figure A.5).

We need an instrument for culture-destination migration flows that is uncorrelated with cultural amenities. We use distance and coal deposits to predict migration flows to destinations and thus predict how shocks to cultural choices in origin locations propagate through to destination populations. Finally, we recenter the instrument following Borusyak and Hull (2020). We discuss the construction of this instrument in Appendix D.2. The key assumption for identification is that this instrument is uncorrelated with factors distinct from the destination population that attract members of a culture to a given location. While we cannot directly verify that assumption, we can verify that the instrument is uncorrelated with observable measures of cultural similarity between locations. We discuss this check in Appendix D.3.

We estimate α by two-stage least squares with fixed effects for destinations and for the

interaction of the destination culture (based on the surname-based cluster it is allocated to) and the culture in question. This latter set of fixed effects account for the concern that the surname cluster a district is in might be correlated with cultural amenities.

Given an estimate of α, we can then regress $\gamma_d^k - \alpha \ln m_d^k$ against destination fixed effects to recover v_d^θ (the exponential of the destination fixed effect) and $\left(\xi_d^k\right)^\theta$ (the exponential of the residual).

5.3 *Estimating Cultural Choice Parameters*

In the third stage we estimate the parameters linking cultural choices to migration opportunities. Taking estimates of inverse migration costs, δ_{od}^θ, and the bundle of factors that attract people to a given location, $\left(v_d \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha$, from 5.1, we can calculate Ω_o^k, the expected utility in culture k of one born in o, according to equation (2).

Taking logarithms of equation (3) gives the following regression equation:

$$\underbrace{\ln \sigma_o^k}_{\ln \text{share assigned } k \text{ in } o} = \varphi \ln \Omega_o^k - \underbrace{\ln \left(\sum_{l=1}^{K} \left(\Omega_o^l\right)^\varphi \psi_o^l \right)}_{\text{place FE}} + \underbrace{\ln \psi_o^k}_{\text{Error}}. \tag{5}$$

This equation implies that in a regression of the log of the share choosing culture k in location o (σ_o^k) against the log of Ω_o^k and a location fixed effect, the coefficient on $\ln \Omega_o^k$ corresponds to φ, and the residual corresponds to the logarithm of ψ_o^k.

Estimating (5) is complicated by endogeneity. An exogenous preference for choosing a given culture in a given location—ψ_o^k, the error term in equation (5)—causes there to be more people of that culture resident in proximate locations. As Ω_o^k is increasing in the number of members of culture k migrating near o, a local increase in the propensity to choose culture k should increase Ω_o^k, the independent variable. A secondary source of endogeneity is that cultural transmission tastes ψ_o^k may be correlated with cultural amenities ξ_o^k, which also features in the independent variable Ω_o^k. Places with strong Anglican cultures as measured by church attendance had stronger transmission tastes for the Southeast culture, but also were

more attractive to migrants from the Southeast culture (Figure A.5).

As in the previous section, we use the interaction between coal deposits encouraging migration, and the spatial distribution of historical cultures, to develop an instrument. We first predict the migration attraction of each location using coal deposits. We then allocate each destination location to a culture based on the surname cluster it belongs to. Finally, we average across locations within a cultural cluster to predict the migration pull of each culture. Appendix D.1 provides further detail. We recenter the instrument following Borusyak and Hull (2020). Appendix D.3 confirms that this instrument is also uncorrelated with observed cultural similarity. We use Poisson Pseudo-Maximum Likelihood to estimate φ because it can account for zeros in the cultural choice share (σ_o) and fits the multinomial logit structure of the cultural choice (Baker, 1994).[10] As recommended by Wooldridge (2010), we use the control function approach for instrumental variables estimation. We first run an OLS regression of $\ln \Omega_o^k$ against the instrument and fixed effects, and include the residuals from this first stage regression in the second stage. We include fixed effects for origins and for the combinations of the surname culture each location is allocated to and the culture in question, as above. Given an estimate of φ, we can back out ψ_o^k.

6 RESULTS

In this section, we present our main results. We begin by describing our estimates, with an emphasis on the two key structural parameters: the homophily and culture elasticities. As validation, we show that additional estimated taste and economic parameters correlate with observable proxies and thus appear to be capturing something meaningful. The model's prediction based on the difference between 1911 cultural choices and a counterfactual estimated under 1851 parameters strongly correlates with observed cultural change during this period, offering further support for the model's validity. Importantly, the model reproduces the main

[10] To see how Equation (3) corresponds to multinomial logit, note that it can be rewritten as $\sigma_o^k = \frac{\exp(\varphi \ln \Omega_o^k + \ln \psi_o^k)}{\sum_{l=1}^{K} \exp(\varphi \ln \Omega_o^l + \ln \psi_o^l)}$

spatial patterns we observe in the data: the rise of the Southeast culture and core-periphery differences in identity change.

We next use the model and conduct a set of counterfactual exercises to gain additional intuition on mechanisms and answer the questions we posed at the start of the paper: Which culture emerges as focal in the absence of centralized nation-building? When does industrialization lead to cultural change or retention? We also use the counterfactual exercises to clarify the contribution of different components of modernization to identity change, showing that peripheral homogenization is mainly due to falling migration costs, while industrial development has heterogeneous effects on identity change.

6.1 *Parameter estimates and validation*

ESTIMATION RESULTS Table 3 shows the results of the estimation routines for α and φ. Given that the bias from endogeneity on the non-instrumented coefficients should be positive—in both cases the regressor should be positively correlated with the error term—it is encouraging that the instrumented coefficients (models 3 and 6) are smaller than the non-instrumented coefficients (2 and 5).[11]

We estimate a homophily elasticity α equal to 0.52. The fact that migration choice is responsive to the number of people from the same culture present in a location accords with a large empirical literature on migrant networks which shows that the presence of co-ethnics or co-culturalists is highly predictive of migration decisions (Munshi, 2020; Card, 2009; Burchardi, Chaney and Hassan, 2019). More novel to our setting, we also estimate a positive culture elasticity. Our instrumental variables estimate implies that $\varphi = 1.80$.[12]

[11]We report conventional robust standard errors clustered by district in models (1)–(5) of Table 3 and cluster-bootstrap (6). In doing so we follow other papers that estimate quantitative spatial models in multiple stages (see for instance Donaldson and Hornbeck 2016; Donaldson 2018; Bryan and Morten 2019; Fajgelbaum and Redding 2022; Morten and Oliveira 2023 and Allen and Donaldson 2020). A limitation of conventional standard errors is that they do not take into account uncertainty from previous stages of the estimation. In Appendix E Table E.1 we report confidence intervals from bootstrapping the entire estimation procedure for the model elasticities. We note however that inference in quantitative spatial models is not straightforward because of dependence between all observations; to the best of our knowledge there is no scholarship on how to correctly account for such dependence.

[12]Results from Allen, Arkolakis and Li (2020) establish that a sufficient—though not necessary—condition

	$\ln m_d^k$	γ_d^k		$\ln \Omega_o^k$	$\ln \sigma_o^k$	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	0.641					
	(0.172)					
$\ln m_d^k$ $[\alpha]$		0.604	0.521			
		(0.018)	(0.051)			
Recentered coal-predicted $\ln \Omega_o^k$				0.214		
				(0.035)		
$\ln \Omega_o^k$ $[\varphi]$					2.185	1.803
					(0.103)	(0.304)
First-stage residuals						0.388
						(0.285)
Model	First stage	OLS	TSLS	First stage	PPML	Control function
First stage F-stat			13.867			38.416
N	7418	7418	7418	7452	7452	7452
R^2	0.989	0.999	0.999	0.998		

This table shows estimates of α and φ. All models are estimated at the district-by-culture level. Models (1) and (4) show the first stage, (2) and (5) the non-instrumented second stage and (3) and (6) the instrumented second stage. (1), (2), and (4) are estimated by OLS, (3) by TSLS, and (5) and (6) by Poisson Pseudo-Maximum Likelihood, with (6) including the residuals from (4) as a control function. In models (2) and (3) the dependent variable is the destination-by-culture fixed effect from an origin-by-destination-by-culture gravity model, in (5) and (6) it is the log share of people born in the district assigned to the culture. All models include fixed effects for the district and the culture interacted with the historic culture of the district. Models (1)–(3) are weighted by the number of individuals at the destination, (4)–(6) by the number at the origin. Standard errors clustered by district in parentheses—in (6) these are by bootstrapping (4) and (6) with a fractional random weight bootstrap clustered by district.

Table 3: Estimates of model elasticities

The finding that parents consider the expected value of a culture for their children when making cultural choices is not ex ante obvious, as many studies emphasize parents' preferences to preserve their own cultures in socialization decisions (see examples discussed in Bisin and Verdier (2011)). At the same time, our estimates are consistent with research which shows that parents are responsive to the environment, such as economic opportunities for their children, when making cultural investments, including the choice of first names (Algan et al., 2022). A positive φ also implies a force that mitigates the extent to which cultural diversity increases as a result of migration. This is of more general interest as it suggests that anticipated migration—such as that driven by economic opportunity—should result in more homogeneity than forms of migration that are driven by external shocks, such as forced displacement. In the context of internal labor migration, a positive culture elasticity pushes towards a more homogeneous cultural map. Finally, in the context of the framework of Bisin and Verdier (2001), the estimated elasticities jointly point to strategic complementarities in parental choice of culture. While studies often emphasize parents' preferences for their own cultures as a driver of cultural persistence, when parents' cultural choice depends on the economic and social environment they expect their children to face, parents may choose to socialize their children in cultures different from their own (Bisin and Verdier, 2011, pp.359-362). Our estimates suggest that, when homophily influences migration decisions, parents are more likely to choose cultures held by more people. Strategic complements, in turn, generate coordination dynamics in the emergence of a focal identity, consistent with the work of Laitin (1994) on language.

PARAMETER VALIDATION The model produces estimates of the attractiveness to migrants of each location (v_d^θ), and of factors that made locations attractive to migrants of specific cultures (ξ_d^k), and that made parents more likely to choose specific cultures in a given location

for equilibrium uniqueness in our model is $\frac{\alpha}{1-\alpha}\max(2\varphi-1,1)<1$. We present the derivation of that result in Appendix C. That condition is not satisfied for these estimated elasticities. In practice, we find that regardless of starting values—including starting values that place almost all weight on one culture—the model converges to the same equilibrium with these elasticities.

(ψ_o^k). To validate the location-culture taste parameters, ξ_d^k and ψ_o^k, we examine whether they correlate with the district's religious and surname similarity to districts assigned to the culture in question. We calculate Euclidean distance between the shares of worshippers in different denominations across districts in the 1851 Census of Religious Worship (Southall and Ell, 2022). Similarity to the culture in question should influence cultural and migration choices. We indeed find this to be the case in Figures 5 and 6.

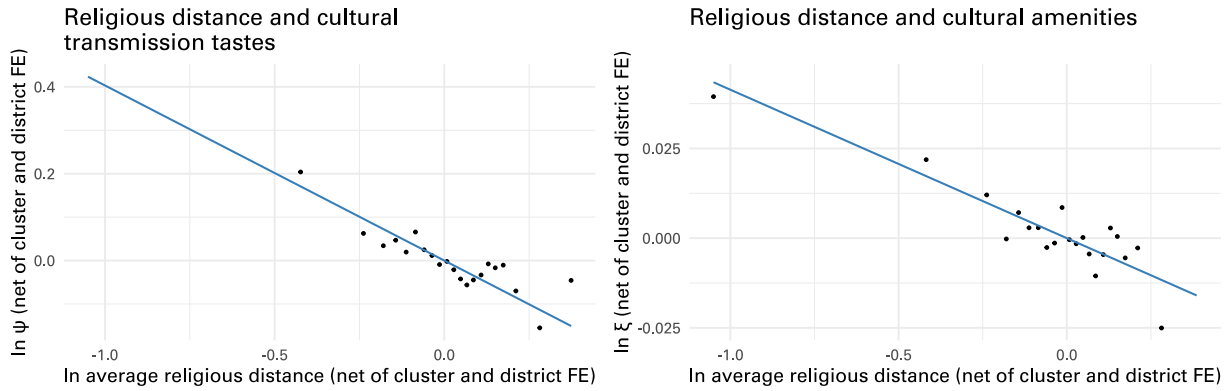


Figure 5: Binned scatterplots of cultural transmission tastes ψ_o^k, cultural amenities ξ_d^k and the average Euclidean distance of religious denomination shares in a district to those in a given cluster in 1851. The figures residualize out cluster and district fixed effects.

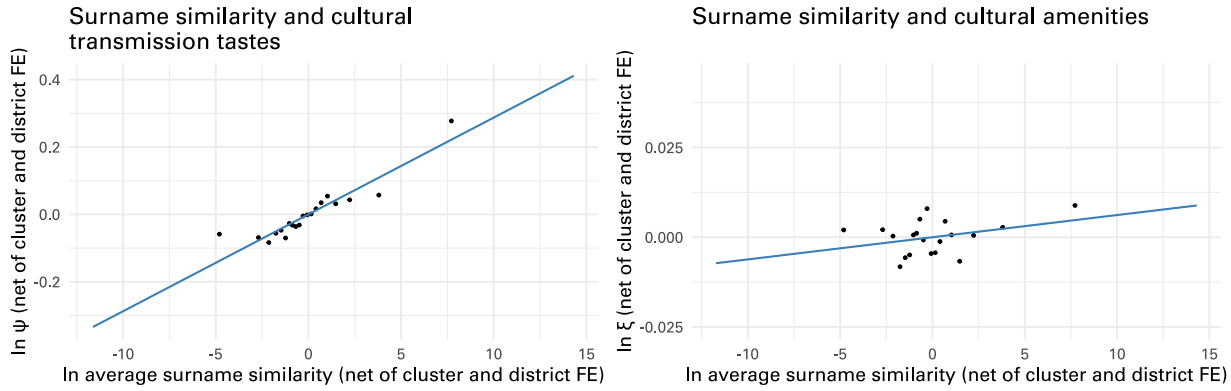


Figure 6: Binned scatterplots of cultural transmission tastes ψ, cultural amenities ξ and the average similarity of surnames in a district to those in a given cluster in 1800

To validate the measure of real wages, v_d^θ, we examine its correlation with agricultural and

manufacturing employment. In a period of structural transformation away from agriculture, we would expect the highest wage places to have less employment in agriculture, and more in manufacturing. Figure 7 shows that this is the case.

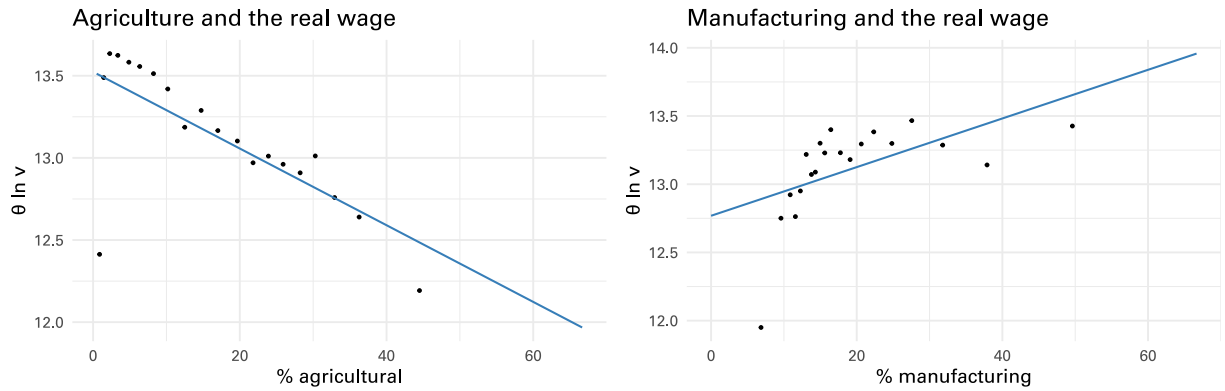


Figure 7: Binned scatterplots of shares of employment in agriculture and manufacturing and real wages $\ln v_d^\theta$

OUT OF SAMPLE VALIDATION As additional validation, we show that the model predicts which cultures grow in a given district, and which decline. We solve for equilibrium cultural choices given 1851 real wages, migration costs, and starting populations, and 1911 cultural amenities and transmission tastes. One can think of the difference between cultural choice shares from this simulation, and 1911 observed cultural choice shares, as the model's prediction of how economic changes over the 1851–1911 period influenced cultural changes, holding fixed noneconomic factors influencing culture, which are embedded in the cultural amenity and transmission taste terms, ξ_d^k and ψ_o^k. Holding these ξ_d^k and ψ_o^k terms fixed accounts for the many other developments in this period, for instance in education, that affected cultural choices through channels other than migration.[13]

[13]We estimate the economic components of the model using data on district-to-district lifetime migration from the 1851 Census. Our estimation routine follows that in Section 5.1, using observed migration flows to estimate pull factors and distance-related costs. The key difference relative to 5.1 is that we do not have naming data for the 1851 period and so cannot use origin-destination-culture migration flows. In practice, this means that we estimate real wages and migration costs—variables that do not vary across cultures—and need to rescale our estimates of real wages to account for the component of utility that is increasing in the number of members of a given culture. Appendix F describes the estimation in greater detail. We also use

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.195	0.173	0.182	0.115
	(0.012)	(0.012)	(0.017)	(0.016)
District FE		x		x
Cluster FE			x	x
N	7362	7362	7362	7362
R^2	0.044	0.099	0.533	0.589

This table shows OLS estimates at the district-by-cluster level. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table 4: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort

In Table 4 we compare these predicted changes in cultural shares against observed changes in cultural shares between those born 1841–1860 and those born 1861–1895. The table reports the results of district-cluster level regressions in which the dependent variable is the change in the log share choosing each culture between the 1841–1860 and 1861–1895 generations, and the independent variable is the observed log share from 1861–1895 minus the log share predicted by our counterfactual using 1851 economic fundamentals. There is a strong positive correlation between the observed and counterfactual changes, that is robust to the addition of district and cluster fixed effects. The latter is important, as it indicates that the model successfully predicts more subtle within-cluster variation, rather than just predicting which clusters experienced increases in popularity.

Our model not only matches the observed change on average, but also reproduces the more subtle spatial patterns of observed change. First, the model captures the increase in the popularity of the culture of the Southeast of England, which is especially strong on the periphery. The increase in the Southeast culture can be seen in the first row of Table 5, which presents changes in the popularity of the home and Southeast cultures, and migration rates,

the starting populations of migrants in 1851 to provide an estimate of starting populations.

under different counterfactual scenarios. Setting the model's exogenous fundamentals—real wages, migration costs and starting populations—to 1851 levels decreases the prevalence of the Southeast culture by around 30%. The center-periphery gradient of this change is illustrated in Figure 8A, which plots the observed log share for the Southeast culture minus the counterfactual log share against distance from London. This predicted change in the popularity of the Southeast culture is large and similar in magnitude to what we see in the data, shown in Figure 8D.

This change helps put our estimated elasticities into perspective. In this counterfactual, namely keeping wages and migration costs at their 1851 levels, the average location experiences a 0.12 log unit decrease in the Southeast population residing there, which corresponds to a 6% reduction ($\exp(-0.12 \cdot \alpha) - 1 = -0.059$) in the attractiveness of that location to migrants from the Southeast. The average location experiences a 0.25 log unit decrease in expected migrant utility under the Southeast culture (Ω_o^k) relative to other cultures, which corresponds to a 36% decrease ($\exp(-0.25 \cdot \varphi) - 1 = -0.36$) in the weight parents attach to the Southeast culture when choosing which culture to assign their children.

Second, the model recovers the heterogeneous effect of coal for home culture loss observed in Table 1. This is shown in Figure 9A. Relative to other peripheral districts, peripheries with coal retain their home cultures. While our model abstracts from many developments of the period that ought to contribute to cultural change, this exercise suggests that the mechanisms we do focus on—economic change driving migration and cultural sorting—plausibly account for the specific patterns of identity change we observe in our data.

6.2 *Mechanisms behind the patterns of identity change*

To gain more intuition into the mechanisms driving these patterns, we conduct two additional counterfactual exercises. Two distinct facets of Britain's modernization during the Second Industrial Revolution enter our model as economic fundamentals. One is industrialization (Figure 7), which corresponds to the the real wage in different locations, v_d^θ and influences

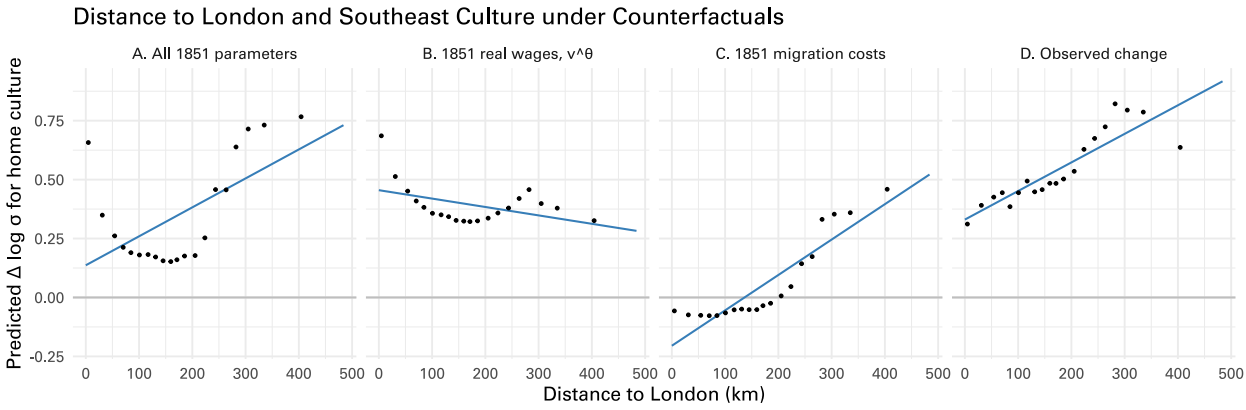


Figure 8: Distance to London, coal, and the changing popularity of the Southeast English culture over the late 19th century

This figure shows the predicted change in the log share allocated names most associated with the Southeast culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance from the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B, only 1851 real wages, and C, 1851 migration costs. Panel D shows the observed change going from the 1841–1860 generation to the 1860–1895 generation.

migration decisions. The second one are the costs of migration, which also change over time. In each counterfactual, we fix other economic fundamentals—including cultural parameters and populations—and solve for the equilibrium of the model under 1851 real wages or migration costs. Effectively, we compare England and Wales in 1911 to a hypothetical country where everything is the same, but real wages or migration costs, respectively, remained at their 1851 levels. The second and third row of Table 5 present the results of these counterfactual exercises. By studying the contribution of each form of economic change separately, we disentangle their effects and use the model to understand which culture will emerge as focal and whether industrialization always leads to cultural change.[14]

FOCAL CULTURE Our model predicts an increase in the popularity of the SE. Why does the London culture win out in particular? In classic accounts of nation-building, the state imposes the national culture through education and conscription (Weber, 1976). In the

[14]In Appendix H we separately examine the effects of changing origin populations, which operates similarly to changing real wages. Increasing the origin population in a location increases the stock of available migrants assigned to the cultures popular in that location, which in turn enters into others' migration and cultural choices. Population growth in industrializing regions reinforces the effects of wage growth.

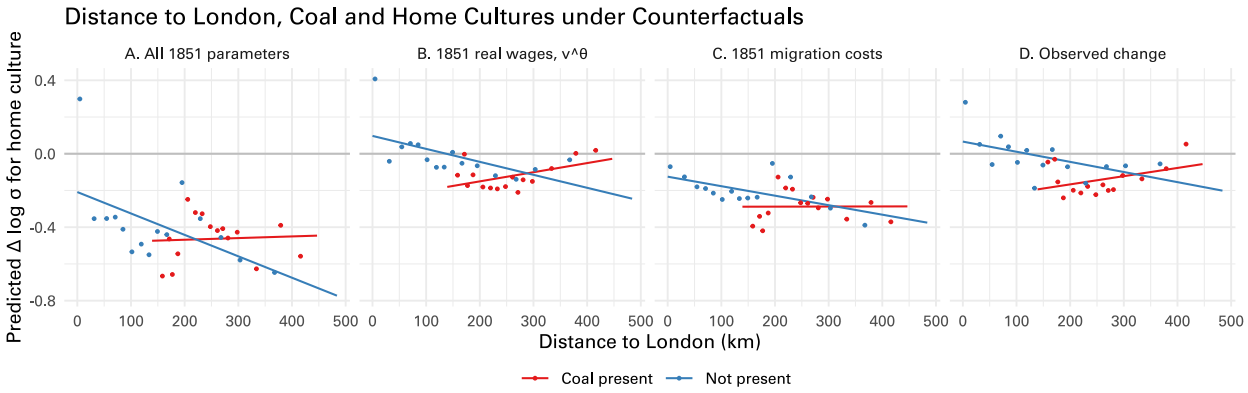


Figure 9: Distance to London, coal, and the changing popularity of home cultures over the late 19th century

This figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance from the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B, only 1851 real wages, and C, 1851 migration costs. Panel D shows the observed change going from the 1841–1860 generation to the 1860–1895 generation.

context of the model, such policies would be captured in the transmission taste (ψ_o^k) terms. In the counterfactual analyses we hold these terms fixed, in order to understand how bottom-up mechanisms of identity choice contribute to the rise of a national culture.

Our results indicate that the emergence of the Southeast as the focal culture was mainly driven by the spatial distribution of industrial change. The second row of Table 5 shows that, had wages remained in their 1851 level, the popularity of the Southeast culture would be 35% lower, a decrease larger than the entire decrease predicted by the counterfactual using all model fundamentals. The predicted change in cultural shares from this real wage counterfactual explains around a quarter of the raw variation in observed changes, much of which is attributable to variation in which cultures grew and declined (Table B.5).

Why does this occur? In the model, cultural choice depends on the migration opportunities (Ω_o^k) faced by members of each culture in each location, which are given by the sum over destinations of real wages (v_d^θ), destination populations (m_d^k) and cultural amenities (ξ_d^k), weighted by migration costs (δ_{od}^θ). Cultural choice follows changes in the components of Ω_o^k. Specifically, parents choose cultures that are likely to benefit their children in locations

Table 5: Counterfactual estimates

	Home culture pop.		S E Culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	3.193	-0.611	-31.448	-28.323	-35.670	-30.983
1851 real wages, v_d^θ	-16.692	-16.692	-35.070	-35.070	-1.562	-1.562
1851 migration costs	16.117	16.117	-1.058	-1.058	-31.286	-31.286

This table shows percentage changes in the population choosing each district's home culture, the population choosing the Southeast English culture, and the share migrating from their district of birth. In each group, the first column gives the total percentage change in the quantity of interest under the counterfactual relative to the observed outcome, the second gives the percentage change implied by the share in each location choosing it, fixing origin populations at the observed level. To fix populations, we calculate the weighted average of the quantity of interest in the counterfactual over locations weighted by realized starting population, and dividing that by the weighted average of the realized quantity of interest, weighted by realized starting population. The first row replaces destination real wages, migration costs and starting populations with their estimated 1851 values, the second replaces only destination real wages, the third replaces migration costs with those estimated from 1851 data.

that become likelier migration destinations due to economic development and rising real wages. When a location d grows, parents become more likely to choose cultures that have large populations in d (because their children will prefer to reside with members of the same culture) and that have valuable amenities in d. Furthermore, parents will weigh these factors more for destinations that are closer to them, as those will be more relevant to their children's migration choices. Finally, cultural choices depend on the relative appeal of different cultures. Wage growth in a destination will increase the popularity of a given culture if the destination specifically draws migrants of that culture.

Concretely, the model implies that the culture of the Southeast grew because economic development during the 1851–1911 period primarily occurred in locations that were well-positioned to attract migrants of the Southeast culture or with high amenities for the Southeast culture *and* because that growth was not offset by development in other regions of the country that would have favored different cultures.

We show this by calculating the model's prediction of how wage growth in each region affects the national popularity of each culture (see Appendix G for details). This culture-wage elasticity captures the responsiveness of each culture to economic development in a location. The measure takes into account the amenities and destination populations that draw members of specific cultures to each location, how easily each location can attract migrants, and how

far it specifically attracts members of one culture relative to others, which magnifies its influence on origin cultural choices.

Figure G.1 shows that the Southeast culture would grow with economic development in the area around London, and would shrink with wage growth in Lancashire, South Wales, or the Northeast. Figure G.2 shows that actual wage growth between 1851 and 1911 most positively correlates with the culture-wage elasticity for the SE. This means that wage growth occurred in places where it is predicted by the model to benefit the Southeast relative to other cultures; in practice, in the region around London and not in Lancashire, the center of the First Industrial Revolution.

INDUSTRIALIZATION AND HETEROGENEITY IN IDENTITY CHANGE. Does industrialization always lead to cultural loss and the adoption of new identities, as most scholars of modernization and nationalism have claimed (Marx and Engels, 1978 [1848]; Deutsch, 1966; Gellner, 2006 [1983]; Green, 2019)? The counterfactuals point to a different conclusion.

Changing industrial activity leads to a heterogeneous decline of local cultures. Figure 9B shows that in the real wage counterfactual, locations with coal experience a decline in the popularity of their home cultures, but only if they are close to the center. Distance from London is positively associated with home culture retention for coal-producing locations. The opposite pattern is true for non-coal-producing ones. This pattern closely matches the observed data (reproduced in Figure 9D for ease of comparison). Table B.6 shows that—in both the observed data and model output—wage growth was associated with the growth of the home culture, but only in districts sufficiently far from London.

Appendix I directly studies the heterogeneous relationship between wage growth and the popularity of different cultures. For each location, we run a separate simulation, examining how raising wages 1% changes cultural choices. That exercise differs from what we have been doing so far, in looking at a hypothetical shock to each location, rather than the observed wage changes to all locations. In these simulations, as one goes further from London, wage

growth also has more positive effects on the popularity of the home culture.

The heterogeneous impact of wage growth is due to migration patterns. Industrial development in a location reduces incentives to emigrate and so increases the importance of cultures that are attractive in that location for identity choice. Local cultures are favored when this mechanism is strong enough to offset any resulting in-migration from other cultures. This is what happens in peripheral locations: because there is less migration to and from the periphery, there are fewer members of other cultures, which makes it less appealing to identify with a non-home culture if one remains in the periphery (see Figure A.6 on the role of distance constraining migration to and from the periphery).[15] Figure A.11B and Table B.7 show that in the counterfactual setting wages to 1851 levels, wage growth decreases out-migration, especially on the periphery.

MIGRATION COSTS AND PERIPHERAL HOMOGENIZATION. If industrialization helps preserve local identities in the peripheries, what contributes to the pattern of cultural loss with a core-periphery gradient observed in the data? Table 5 and Figure 8C show that the reason behind the overall decline in local cultures and greater decline in districts further from the center was falling migration costs, and not industrial development. Though the process of modernization and structural transformation commonly involves changes in both these dimensions, our model helps clarify their distinct effects on cultural change.

The third row of Table 5 shows that keeping migration costs at their 1851 level, holding all else fixed, results in a 16% increase in the share of the population that adopts the home culture. This change is driven by migration: 1851 migration costs imply a 31% decrease in the share of the population migrating outside the district of birth. Lower migration costs imply that more distant destinations (and cultures) weigh more in the cultural choice of parents, driving an overall decline in the prevalence of home cultures.

Lower migration costs imply more migration over longer distances: migration to and from

[15] Which cultures are attractive in a region depends on endogenous migration (m_d^k) and exogenous tastes (ξ_d^k). Peripheral regions have home cultures that are strong due to the endogenous component, population, not taste (see Figures I.3 and I.4, which decompose the contributions of population and cultural amenities).

the Southeast increases most in places further from the Southeast (Figures A.9C and A.10C). The same mechanism accounts for the larger observed declines in local cultures in districts further from the country's center (Figure 9C). Peripheral districts experience a bigger change than central ones in terms of how much they weigh distant cultures. This core-periphery pattern is similar in locations with and without coal. Lower migration costs increase the incentive to migrate from both high and low-wage locations and therefore contribute little to cultural persistence in industrializing peripheries.

While changes in migration costs drive an overall decline in local cultures and more so in districts further from the center, they do not favor the overall spread of the Southeast culture. Lower migration costs increase migration from peripheral regions to the Southeast, increasing the prevalence of the Southeast culture in the periphery, but also increase migration from the Southeast to other regions, decreasing the prevalence of the Southeast culture in the Southeast. In short, our results indicate distinct contributions of different components of modernization to cultural change: industrial development is what determines *which* cultures will grow in popularity, while migration costs determine *where* that will happen.

7 SUMMARY OF ADDITIONAL RESULTS AND ROBUSTNESS CHECKS

This section reviews a number of additional analyses that extend or reinforce our findings. First, in Appendix J.1 we examine how cultural homogenization influences economic output. Total homogenization into the culture of the Southeast, calculated by setting the cultural transmission tastes of all other cultures to zero, increases the average v_d^θ by 0.8%. Cultural amenities (ξ_d^k) for the Southeast culture are correlated with higher real wages; assimilation into the Southeast culture moves workers out of low-wage regions with high cultural amenities for other cultures. The small magnitude of this effect makes sense because distance-related migration costs keep workers in low-productivity regions, and the counterfactual is relative to a baseline in which wages strongly influence migration decisions.

Second, we address the concern that our measure of migration costs which relies on

geographic distance does not account for additional factors affecting the ease of moving across locations, such as transport links. In Appendix J.2 we develop measures of transportation costs between locations by creating a network database of railways and shipping routes following Donaldson and Hornbeck (2016). While we do find that the implied transportation costs negatively correlate with migration flows, this association is entirely captured by geographic distance. Measures of migration costs incorporating transportation infrastructure would not improve on the distance-related measures we use in our main analyses.

Third, in Appendix J.3 we examine whether our measures of cultural amenities actually capture local cultural wage premia, for instance because of discrimination in the labor market. Using data on the economic status of jobs held by members of different cultures in different locations, we find a null or negative relationship between cultural amenities and status. A negative relationship would be consistent with a negative compensating differential for amenities as in canonical economic geography models (Roback, 1982; Bryan and Morten, 2019), which increases our confidence in our interpretation of ξ_d^k as a cultural amenity.

Fourth, in Appendix J.4, we discuss how accounting for emigration affects our estimation and conclusions. Over the course of the 19th century vast numbers of English and Welsh people emigrated, especially to the US, Canada, Australia, and New Zealand. We show that our theory and evidence related to internal migration is unaffected by the existence of emigration, because our data conditions on not emigrating. The potential for emigration does affect cultural choices, but in our framework it is captured by the location-by-culture cultural transmission tastes. Our counterfactual exercises must be interpreted as holding the ratio of domestic to international migration opportunities fixed.

Fifth, in Appendix J.5 we examine whether changing the homophily (α) and culture (φ) elasticities alters our results. Figures J.3 and J.4 show that the broad stylized facts our model reproduces hold at a wide range of feasible elasticities, although the magnitude of the relationship, for instance, between distance to London and the growth of the Southeast, varies considerably.

Sixth, in Appendix J.6 we explore the robustness of our results to alternative cultural clusters. We show that both the growth of the Southeast and the ability of our model to rationalize its rise across geography hold using either regions used to organize civil defense, or clusters estimated using medieval linguistic data.

8 CONCLUSION

We have examined the role of migration in changing the cultural map of 19th century industrializing England and Wales. By modeling strategic location and identity choices in a spatial equilibrium framework we find that industrialization-induced migration can favor assimilation into a focal identity through coordination dynamics. The central contribution of our study is to highlight the importance of the spatial distribution of economic development for the resulting cultural map. Where industrialization occurs in space determines which identities will become focal. The interaction of industrial development and migration costs, in turn, determines which regions will assimilate into those focal identities. Industrialization promotes assimilation in central locations, but contributes to cultural retention in the peripheries.

Insights from our study elucidate the role of labor migration for cultural change in 19th century Britain and complement existing explanations for the survival of regional differences. Most prominently, Hechter (1977) proposed a theory of economic and sociopolitical dominance of the peripheries by the core to explain the strength of local identity in Wales, Scotland and Ireland. Our model supports some of the patterns and mechanisms identified by Hechter, such as that industrialization drives overall homogenization and that decreased out-migration from Wales in response to coal-driven industrial development promoted the retention of the Welsh language (until English language policies explicitly targeted it for eradication in the early 20th century). As in Hechter, our model also emphasizes heterogeneity within the peripheries. However, while his theory predicts highest assimilation in peripheral industrial hubs, our emphasis on the interplay between labor migration and cultural choice indicates that local industrial development may work in favor, rather than against, the retention of local culture,

counteracting other forces of modernization that pulled the peripheries toward the core.

How do the insights from 19th century Britain travel to other country contexts? Our framework is general enough to apply to any situation in which economic activity changes across space and people make migration decisions driven by both economic and cultural considerations. Empirically, the result that peripheral regions may be more likely to resist assimilation into a dominant culture if they develop economically could rationalize patterns we observe in other parts of Europe during the 19th and early 20th centuries. The Basque country in Spain was able to retain its distinctive identity among other reasons because the development of ore production and steel manufacturing directed local rural migrants to head to Bilbao rather than Madrid (Green, 2022, p.52). Scotland and Catalonia in Spain were additional cases of industrializing peripheral areas with distinct identities. Discouragement of out-migration may have contributed to cultural retention in those regions, while peripherality muted the effects of in-migration from different cultures into their main industrial centers.

The migration link between industrialization and cultural change also informs a broader literature on nation-building. The formation of national identities is thought to be primarily a task of the state, which homogenizes populations through coercion (Tilly, 1975; Mylonas, 2013), conscription (Weber, 1976) or education (Hobsbawm, 1992; Cinnirella and Schueler, 2018; Alesina, Giuliano and Reich, 2021). Our findings show that beyond top-down efforts to build national identities, bottom-up processes such as spatial changes in the pattern of economic activity can also play a crucial role in cultural unification. This in turn has implications for the role of the state, suggesting that less conventional tools such as industrial policy, besides their more well-documented role in restructuring economic activity (Juhász, Lane and Rodrik, Forthcoming), can be used for identity engineering and nation-building by favoring the development of particular economic sectors and locations.

REFERENCES

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122(3):467–506.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm and Nikolaus Wolf. 2015. "The Economics of Density: Evidence From the Berlin Wall." *Econometrica* 83(6):2127–2189.

Alesina, Alberto, Paola Giuliano and Bryony Reich. 2021. "Nation-building and education." *The Economic Journal* 131(638):2273–2303.

Algan, Yann, Clément Malgouyres, Thierry Mayer and Mathias Thoenig. 2022. "The economic incentives of cultural transmission: Spatial evidence from naming patterns across france." *The Economic Journal* 132(642):437–470.

Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics* 129(3):1085–1140.

Allen, Treb, Costas Arkolakis and Xiangliang Li. 2020. On the Equilibrium Properties of Network Models with Heterogeneous Agents. Working Paper 27837 National Bureau of Economic Research.

Allen, Treb and Dave Donaldson. 2020. Persistence and Path Dependence in the Spatial Economy. Working Paper 28059 National Bureau of Economic Research.

Alvarez-Palau, Eduard J., Oliver Dunn, Dan Bogart, Max Satchell and Leigh Shaw-Taylor. 2019. "Historic ports and coastal sailing routes in England and Wales, 1540-1914.". **URL:** *https://reshare.ukdataservice.ac.uk/853711/*

Anderson, Benedict. 1983. *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. New York and London: Verso.

Antràs, Pol and Hans-Joachim Voth. 2003. "Factor prices and productivity growth during the British industrial revolution." *Explorations in Economic History* 40(1):52–77.

Baker, Stuart G. 1994. "The Multinomial-Poisson Transformation." *Journal of the Royal Statistical Society: Series D (The Statistician)* 43(4):495–504.

Bates, Robert H. 1974. "Ethnic competition and modernization in contemporary Africa." *Comparative political studies* 6(4):457–484.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg and Maisy Wong. 2019. "Unity in Diversity? How Intergroup Contact Can Foster Nation Building." *American Economic Review* 109(11):3978–4025.

Bazzi, Samuel, Martin Fiszbein and Mesay Gebresilasse. 2020. "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States." *Econometrica* 88(6):2329–2368.

Bennett, Robert J., Harry Smith, Carry van Lieshout and Gill Newton. 2017. Business sectors, occupations and aggregations of census data 1851-1911. Working Paper 5 University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure. Issue: 5.

Benskin, M., M. Laing, V. Karaiskos and K. Williamson. 2013. "An Electronic Version of A Linguistic Atlas of Late Mediaeval English.".
URL: *http://www.lel.ed.ac.uk/ihd/elalme/elalme.html*

Bisin, Alberto and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97(2):298–319.

Bisin, Alberto and Thierry Verdier. 2011. The economics of cultural transmission and socialization. In *Handbook of Social Economics*. Vol. 1 Amsterdam: North Holland pp. 339–416.

Blanc, Guillaume and Masahiro Kubo. 2024. French. Technical report Working paper.

Borusyak, Kirill and Peter Hull. 2020. Non-Random Exposure to Exogenous Shocks: Theory and Applications. Working Paper 27845 National Bureau of Economic Research.

Bryan, Gharad and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy* 127(5):2229–2268.

Burchardi, Konrad B, Thomas Chaney and Tarek A Hassan. 2019. "Migrants, ancestors, and foreign investments." *The Review of Economic Studies* 86(4):1448–1486.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg and Pierre-Daniel Sarte. 2018. "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy." *The Review of Economic Studies* 85(4):2042–2096.

Caliendo, Lorenzo, Luca David Opromolla, Fernando Parro and Alessandro Sforza. 2021. "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement." *Journal of Political Economy* 129(12):3491–3545.

Card, David. 2009. "Immigration and inequality." *American Economic Review* 99(2):1–21.

Cinnirella, Francesco and Ruth Schueler. 2018. "Nation Building: The Role of Central Spending in Education." *Explorations in Economic History* 67:18–39.

Colley, Linda. 2005. *Britons: Forging the Nation, 1707-1837*. New Haven and London: Yale University Press.

Crouzet, François. 1982. *The Victorian Economy*. London: Methuen and Co.

Day, Joseph. 2016. "Consistent parish cross-walk files.". Work conducted as part of the ESRC-funded project 'Migration, Urbanisation and Socio-Economic Change, England and Wales 1851-1911'.

Day, Joseph. 2023. "Mapping the cultural divides of England and Wales: Did the geographies of 'Belonging'act as a brake on British Urbanisation, 1851–1911?" *Plos one* 18(5):e0286244.

Deutsch, Karl W. 1961. "Social mobilization and political development." *American political science Review* 55(3):493–514.

Deutsch, Karl Wolfgang. 1966. *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality*. Cambridge, MA: MIT Press.

Deutsch, Karl Wolfgang. 1969. *Nationalism and its Alternatives*. New York: Alfred A. Knopf.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108(4-5):899–934.

Donaldson, Dave and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics* 131(2):799–858.

Fajgelbaum, Pablo and Stephen J. Redding. 2022. "Trade, Structural Transformation, and Development: Evidence from Argentina 1869–1914." *Journal of Political Economy* 130(5):1249–1318. Number: 5.

Fally, Thibault. 2015. "Structural gravity and fixed effects." *Journal of International Economics* 97(1):76–85.

Fouka, Vasiliki. 2019. "How do Immigrants Respond to Discrimination? The Case of Germans in the US during World War I." *American Political Science Review* 113(2):405–422.

Fouka, Vasiliki. 2020. "Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I." *The Review of Economic Studies* 87(1):204–239.

Fryer Jr, Roland G and Steven D Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics* 119(3):767–805.

Geary, Frank and Tom Stark. 2015. "Regional GDP in the UK, 1861–1911: new estimates." *The Economic History Review* 68(1):123–144.

Gellner, Ernest. 1964. *Thought and Change*. Chicago: The University of Chicago Press.

Gellner, Ernest. 2006 [1983]. *Nations and Nationalism*. Oxford: Blackwell Publishing.

Green, Elliott. 2019. "Industrialization and ethnic change in the modern world." *Ethnic and Racial Studies* 42(2):178–197.

Green, Elliott. 2022. *Industrialization and Assimilation: Understanding Ethnic Change in the Modern World*. Cambridge University Press.

Görlach, Manfred. 1999. *English in Nineteenth-Century England: An Introduction*. Cambridge: Cambridge University Press.

Hechter, Michael. 1977. *Internal colonialism: The Celtic fringe in British national development, 1536-1966*. University of California Press.

Henneberg, J., M. Satchell, X. You, L. M. W. Shaw-Taylor, E. A. Wrigley and Michael Cobb. 2018. "1881 England, Wales and Scotland Railway Stations.".
URL: *https://reshare.ukdataservice.ac.uk/852996/*

Hobsbawm, Eric J. 1992. *Nations and Nationalism since 1780: Programme, Myth, Reality*. Cambridge University Press.

Homans, George C. 1969. "The Explanation of English Regional Differences." *Past & Present* (42):18–34.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87(5):1439–1474.

John, Christopher R., David Watson, Michael R. Barnes, Costantino Pitzalis and Myles J. Lewis. 2020. "Spectrum: fast density-aware spectral clustering for single and multi-omic data." *Bioinformatics (Oxford, England)* 36(4):1159–1166.

Juhász, Réka, Nathan Lane and Dani Rodrik. Forthcoming. "The New Economics of Industrial Policy." *Annual Review Economics* .

Kandt, Jens, James A Cheshire and Paul A Longley. 2016. "Regional surnames and genetic structure in Great Britain." *Transactions of the Institute of British Geographers* 41(4):554–569.

Laitin, David D. 1994. "The Tower of Babel as a Coordination Game: Political Linguistics in Ghana." *The American Political Science Review* 88(3):622–634.

Laitin, David D. 1995. "Marginality: A Microperspective." *Rationality and Society* 7(1):31–57.

Laitin, David D. 2007. *Nations, States, and Violence*. Oxford, UK: Oxford University Press.

Lambert, Paul S., Richard L. Zijdeman, Marco H. D. Van Leeuwen, Ineke Maas and Kenneth Prandy. 2013. "The Construction of HISCAM: A Stratification Scale Based on Social Interactions for Historical Comparative Research." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46(2):77–89.

Lazear, Edward P. 1999. "Culture and Language." *Journal of political Economy* 107(S6):S95–S126.

Marti-Henneberg, J., M. Satchell, X. You, L. M. W. Shaw-Taylor and E. A. Wrigley. 2018. "1851 England, Wales and Scotland Railway Stations.".
URL: *https://reshare.ukdataservice.ac.uk/852994/*

Marti-Henneberg, J., M. Satchell, X. You, L.M.W. Shaw-Taylor and E.A. Wrigley. 2023. "1881 England, Wales and Scotland Rail Lines.".

Marx, Karl and Friedrich Engels. 1978 [1848]. Manifesto of the Communist Party. In *The Marx-Engels Reader*, ed. Robert Tucker. New York: W. W. Norton pp. 594–618.

McFadden, Daniel. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, ed. Paul Zarembka. New York: Academic Press pp. 105–142.

Mokyr, Joel. 2008. Accounting for the Industrial Revolution. In *The Cambridge Economic History of Modern Britain, Volume 1: Industrialisation, 1700-1860*, ed. Roderick Floud and Paul Johnson. Cambridge: Cambridge University Press pp. 1–27.

Monte, Ferdinando, Stephen J. Redding and Esteban Rossi-Hansberg. 2018. "Commuting, Migration, and Local Employment Elasticities." *American Economic Review* 108(12):3855–3890.

Morten, Melanie and Jacqueline Oliveira. 2023. The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. Working Paper.

Munshi, Kaivan. 2020. "Social networks and migration." *Annual Review of Economics* 12(1):503–524.

Mylonas, Harris. 2013. *The Politics of Nation-Building: Making Co-Nationals, Refugees, and Minorities*. Cambridge University Press.

Palgrave Macmillan Ltd. 2013. *International Historical Statistics*. London: Palgrave Macmillan.
URL: *https://doi.org/10.1057/978-1-137-30568-8*

Pengl, Yannick I, Philip Roessler and Valeria Rueda. 2022. "Cash crops, print technologies, and the politicization of ethnicity in Africa." *American Political Science Review* 116(1):181–199.

Pengl, Yannick, Müller-Crepon Carl Valli Roberto Cederman Lars-Erik and Luc Girardin. Forthcoming. "The Train Wrecks of Modernization: Railway Construction and Nationalist Mobilization in Europe." *American Political Science Review* .

Porteous, J. Douglas. 1982. "Surname Geography: A Study of the Mell Family Name c. 1538-1980." *Transactions of the Institute of British Geographers* 7(4):395–418.

Redding, Stephen J. and Daniel M. Sturm. 2008. "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* 98(5):1766–1797.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6):1257–1278. Number: 6.

Robinson, Amanda Lea. 2014. "National versus ethnic identification in Africa: Modernization, colonial legacy, and the origins of territorial nationalism." *World Politics* 66(4):709–746.

Sambanis, Nicholas and Moses Shayo. 2013. "Social identification and ethnic conflict." *American Political Science Review* 107(2):294–325.

Satchell, A.E.M., P.M.K. Kitson, G.H. Newton, L. Shaw-Taylor and E.A. Wrigley. 2016. "1851 England and Wales Census parishes, townships and places.".

Satchell, M., E.A. Wrigley, L.M.W. Shaw-Taylor, X. You and J. Henneberg. 2023. "1851 England, Wales and Scotland Rail Lines.".

Schürer, Kevin and Edward Higgs. 2014. "Integrated Census Microdata (I-CeM), 1851-1911.". SN: 7481.
URL: *http://doi.org/10.5255/UKDA-SN-7481-1*

Shayo, Moses. 2009. "A model of social identity with an application to political economy: Nation, class, and redistribution." *American Political science review* 103(2):147–174.

Silva, J. M. C. Santos and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4):641–658.

Southall, Humphrey R. and Paul Ell. 2022. "Great Britain Historical Database: Census Data: Religion Statistics, 1851 [data collection] 2nd Edition.".

Tilly, Charles. 1975. Reflections on the History of European State-Making. Princeton, NJ: Princeton University Press pp. 3–83.

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109(5):1843–1872.

Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

Weber, Eugen. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Stanford University Press.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second edition ed. Cambridge MA: MIT Press.

Online Appendix

Table of Contents

A Additional Figures **2**

B Additional Tables **8**

C Equilibrium Uniqueness **16**

D Additional Detail on Instrumental Variables **19**

D.1 φ Equation 19

D.2 α Equation 20

D.3 Validating Instruments 20

E Bootstrapping Elasticities **21**

F Estimating 1851 Economic Fundamentals **23**

G Why the Southeast? **23**

H Counterfactual Effects of Population Change **26**

I Model implied elasticities **28**

J Additional Analyses **34**

J.1 Effects of Cultural Homogeneity on Migration and Economic Outcomes . 34

J.2 Measuring Transportation Costs 35

J.3 Relationship Between Amenities and Wages 36

J.4 Accounting for Emigration 37

J.5 Alternative Elasticities 40

J.6 Alternative Cultural Clusters 44

A ADDITIONAL FIGURES

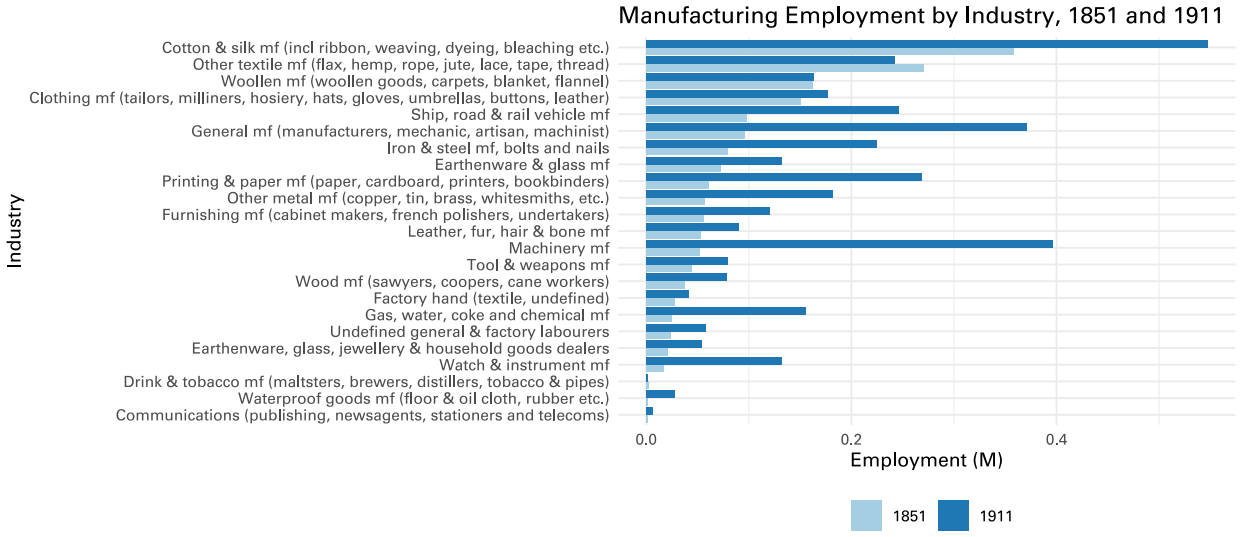


Figure A.1: Employment by manufacturing industry

This figure plots employment in different manufacturing industries in the 1851 and 1911 censuses, calculated using census microdata from Schürer and Higgs (2014) and industry categories from Bennett et al. (2017).

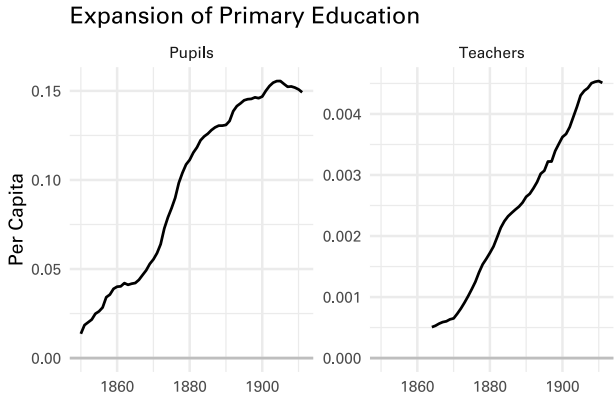


Figure A.2: Primary school teachers and pupils per capita

This figure plots the number of primary school teachers and primary school pupils, divided by total population. Source: Palgrave Macmillan Ltd (2013)



Figure A.3: Map of registration districts highlighting the location of Anglesey

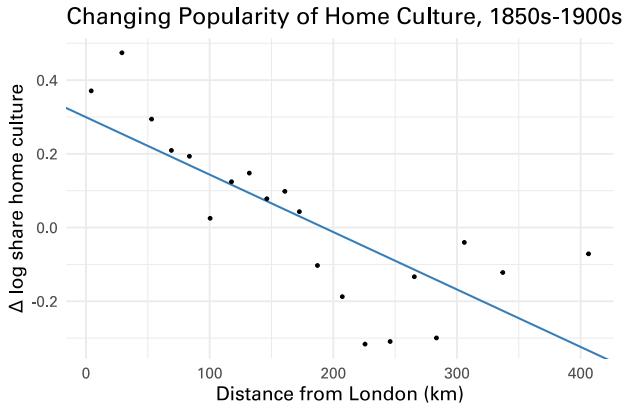


Figure A.4: Change in popularity of the home culture plotted against distance to London

This figure shows the change in the log share allocated names most associated with the home culture comparing those born between 1851–1860 and 1901–1910, plotted against the distance from the district centroid to the City of London.

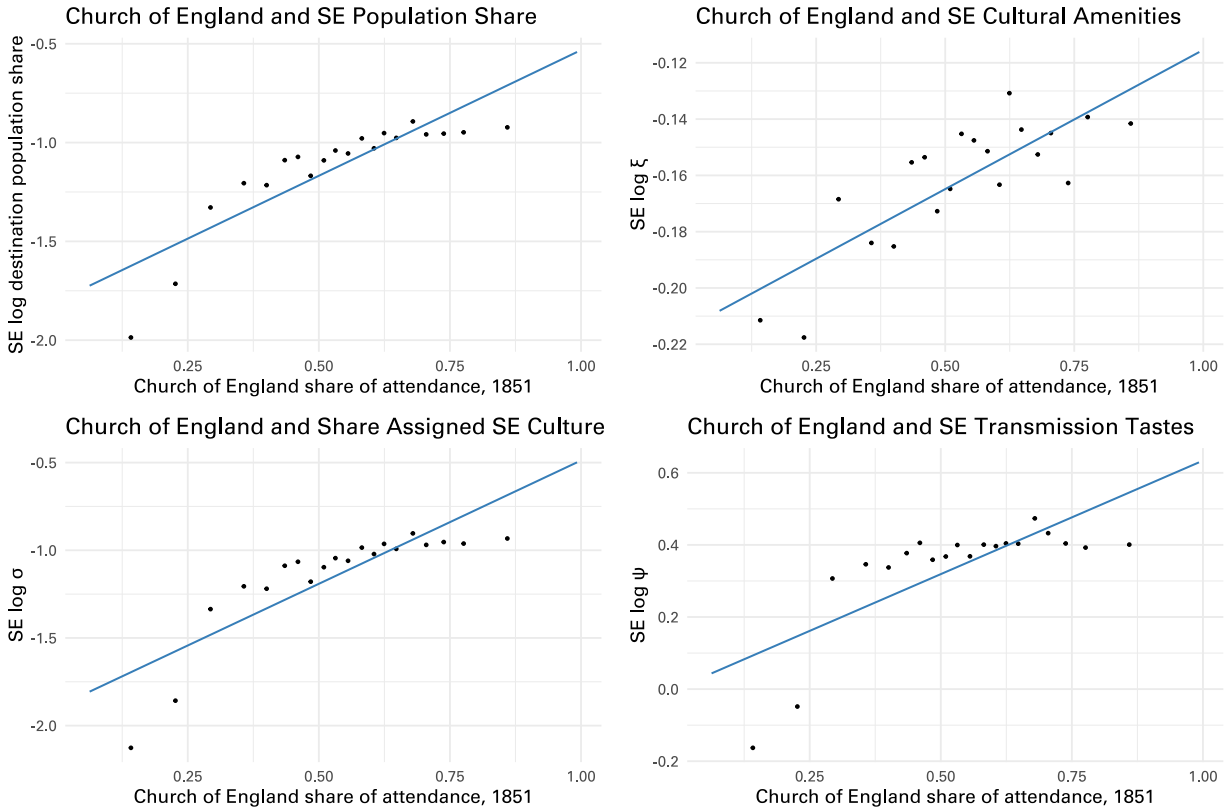


Figure A.5: Places with greater Church of England attendance in 1851 had larger cultural amenities and transmission tastes for the Southeast English culture, as well as origin and destination populations

The top left panel plots the the log share of the destination population assigned to the Southeast English culture, in the 1861–1895 generation, against the Church of England's share of church attendants in the 1851 religious census. The top right panel plots log cultural amenity (ξ_d^k) estimated for the Southeast English culture against that variable. The bottom left panel plots the log share of the origin population assigned to the Southeast English culture against that variable. The bottom right panel plots the log cultural transmission taste (ψ_o^k) for the Southeast English culture against the same variable.

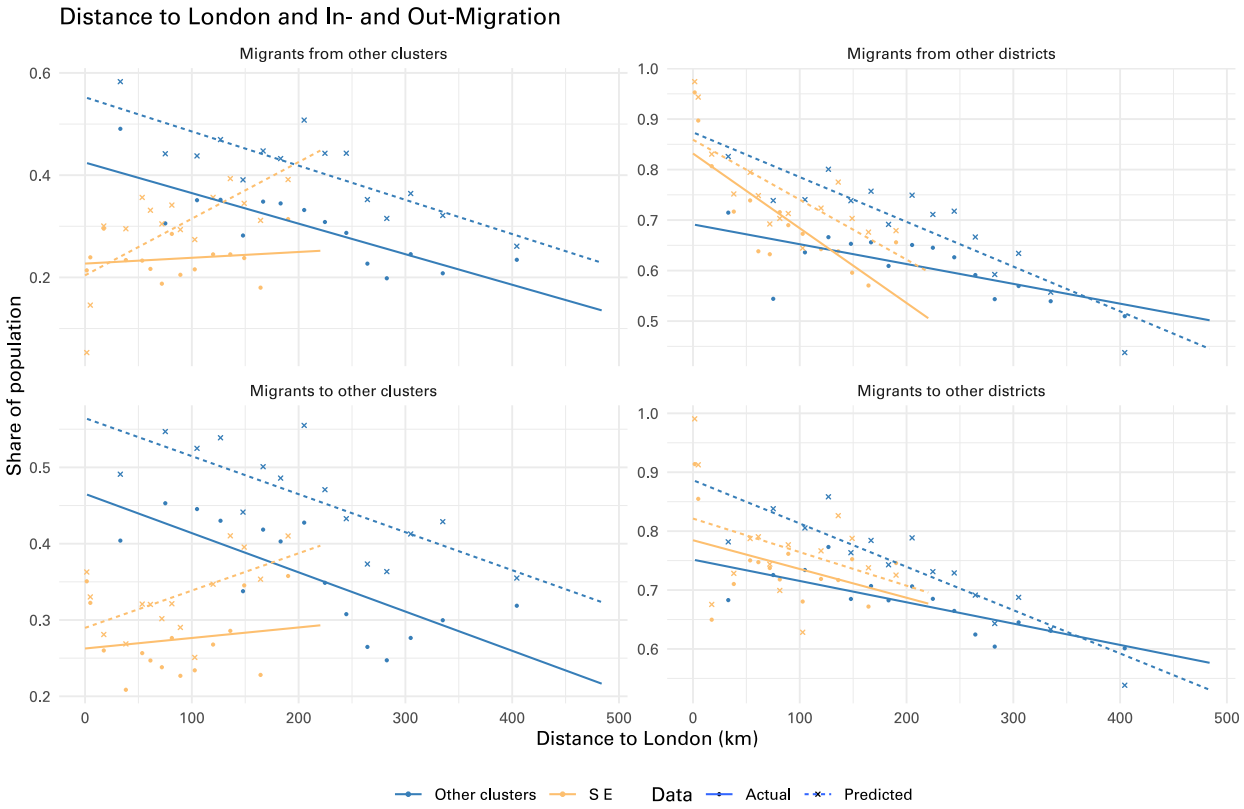


Figure A.6: More peripheral locations had lower rates of in- and out-migration

This figure shows the share of the population migrating into the district from other clusters (top left) or other districts (top right), and the share migrating from the district to other clusters (bottom left), and other districts (bottom right), in the 1861–1895 generation, plotted against geographic distance to London. Data is shown separately for districts allocated to the Southeast culture (in beige), and those allocated to other cultures (in blue). Dots correspond to binned averages, lines correspond to OLS fits. Round dots and solid lines present the observed data. Crosses and dashed lines present predictions from a gravity regression with origin and destination fixed effects and covariates for log geographic distance and an indicator that the origin equals the destination. This regression assumes that geographic location influences migration, but does not assume that people in different locations have different strengths of attachment to their local area. The figure shows that more peripheral locations experience less in- and out-migration at both the district and cluster levels. This pattern is not due to stronger cultural attachments per se, which would not be captured by the gravity regression, but by geographic location, which would be.

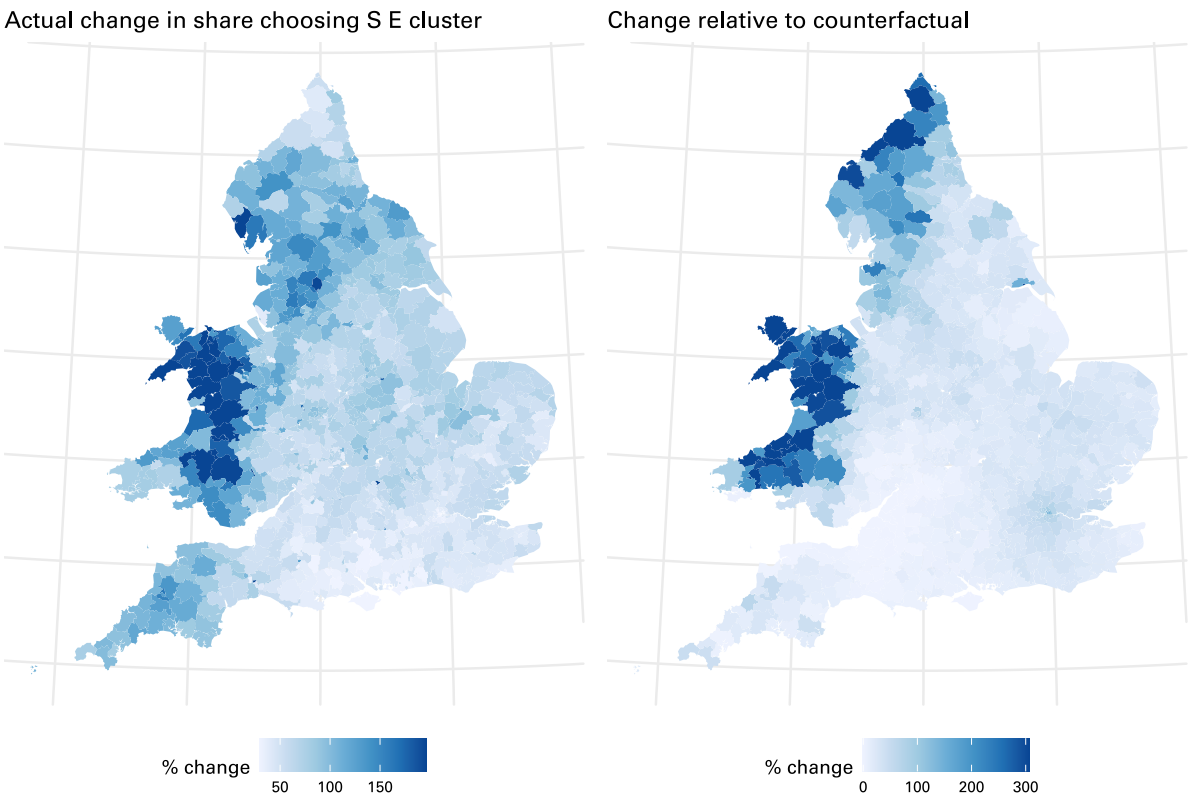


Figure A.7: Change in log share choosing the Southeast England cluster

Left panel is observed change between those born 1841–1860 and those born 1861–1895, right panel is change between counterfactual using 1851 estimates of migration costs, population, and destination utilities, and observed values for those born 1861–1895

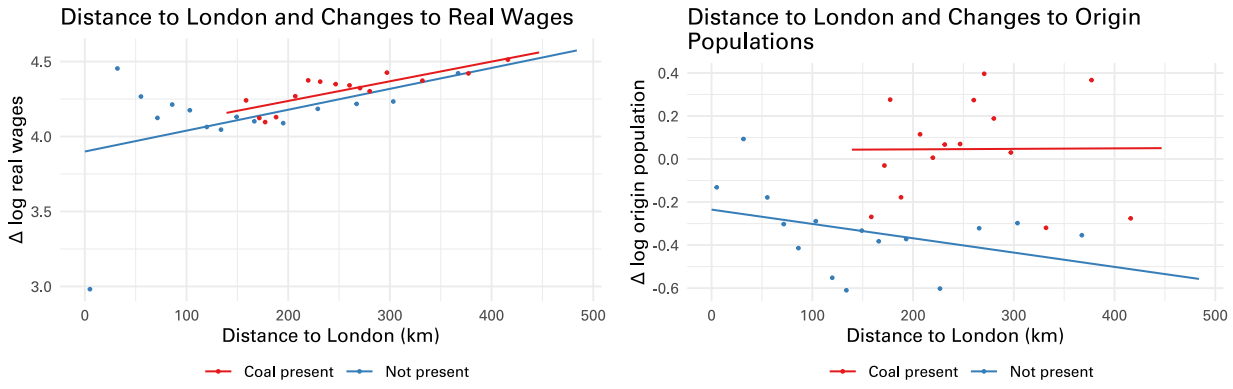


Figure A.8: Change in real wage and origin populations between 1851 and 1911, binned averages plotted against distance from London, subset by whether the district contains coal

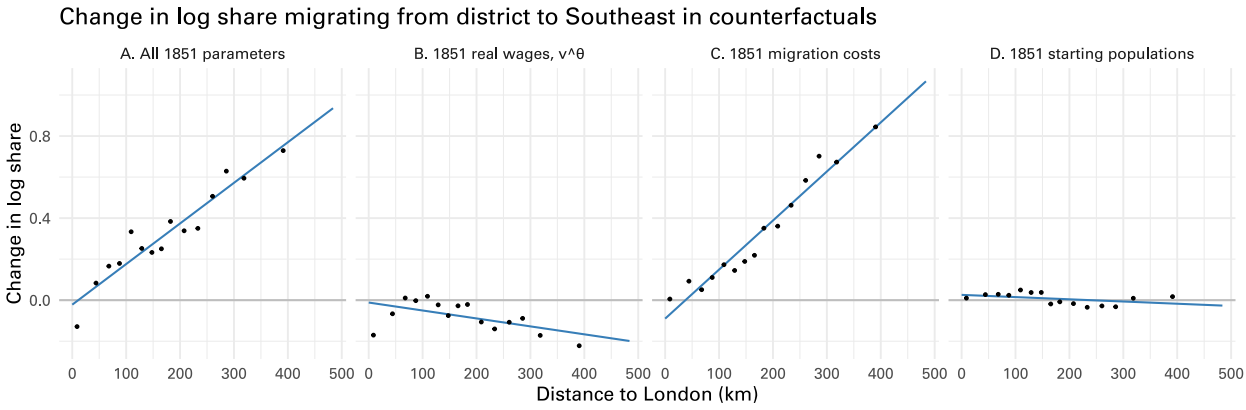


Figure A.9: Counterfactual changes in migration to districts allocated to the Southeast England cluster

The x axis is distance to London, the y axis is observed log share migrating from district in question to districts allocated to the Southeast, minus log share under counterfactual scenario.

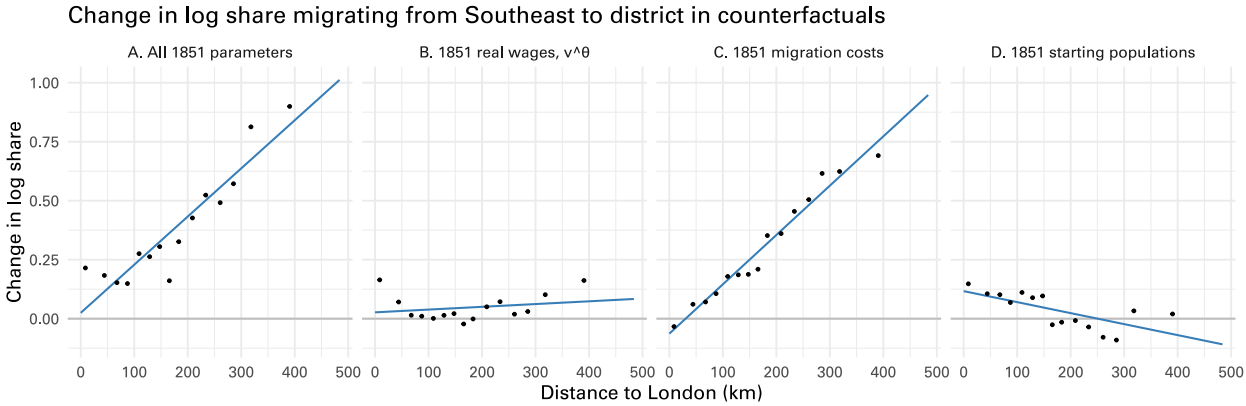


Figure A.10: Counterfactual changes in migration from the Southeast England cluster to districts

The x axis is distance to London, the y axis is observed log share migrating to district in question from districts allocated to the Southeast, minus log share under counterfactual scenario.

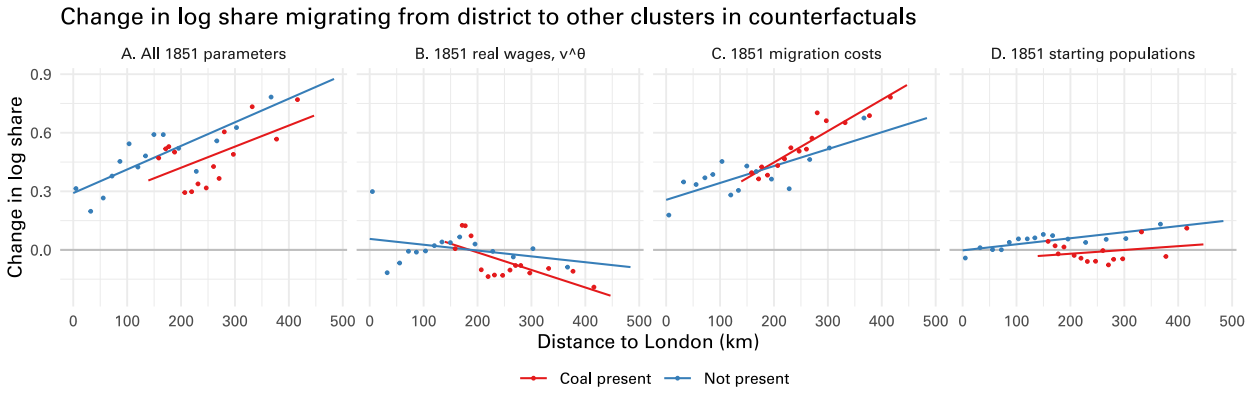


Figure A.11: Counterfactual changes in migration from the district out of the home cluster

The x axis is distance to London, the y axis is observed log share migrating from the district to districts not in the home cluster, minus log share under counterfactual scenario.

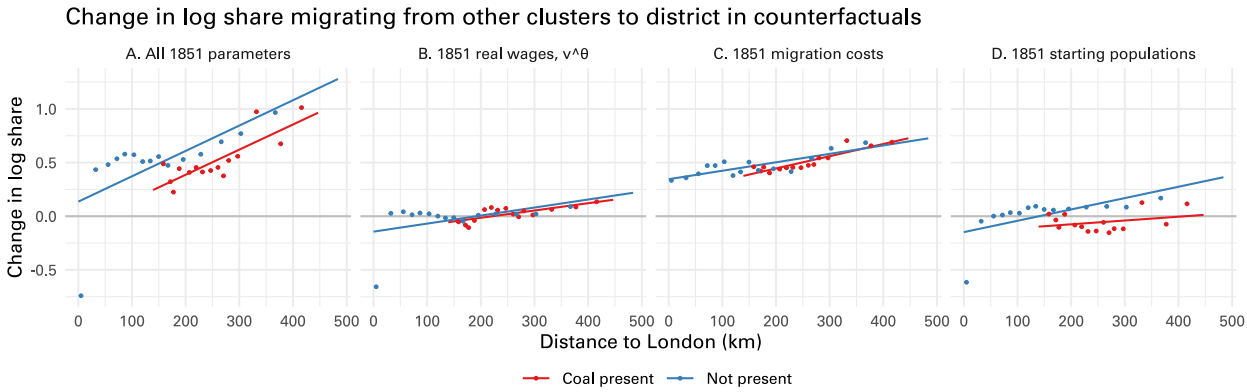


Figure A.12: Counterfactual changes in migration from non-home clusters into district

The x axis is distance to London, the y axis is observed log share migrating to the district from districts not in the home cluster, minus log share under counterfactual scenario.

B ADDITIONAL TABLES

Table B.2: First names with highest name scores by cluster

			Frequency 1841–1860		
Cluster	Metaphone	Example Name	Cluster	Total	Name Score
E	ELFN	Elvina	51	142	0.918
E	HKR	Hagar	25	89	0.886
E	MHL	Mahala	108	449	0.863
E	TMRS	Damaris	17	72	0.860

E	EFRT	Everett	14	67	0.840
E	ETKR	Edgar	241	1190	0.835
E	LRN	Lorna	14	70	0.832
E	OSBRN	Osborne	12	61	0.830
E	BRNBS	Barnabas	12	62	0.827
E	BLNT	Belinda	27	143	0.822
Mid	IX	Isiah	48	86	0.842
Mid	KMFRT	Comfort	37	70	0.823
Mid	RSHN	Rosehannah	53	109	0.798
Mid	IS	Isaiah	218	499	0.767
Mid	SLP	Zilpah	35	81	0.762
Mid	XTRK	Shadrack	53	125	0.755
Mid	SRHN	Sarahann	207	530	0.731
Mid	WLFB	Willoughby	31	79	0.727
Mid	NMN	Newman	29	76	0.726
Mid	TRSL	Drusilla	98	263	0.716
N	TKSN	Dixon	28	63	0.909
N	RXRTSN	Richardson	25	58	0.907
N	JKSN	Jackson	41	107	0.887
N	0MPSN	Thompson	61	162	0.887
N	RBSN	Robson	21	58	0.880
N	WTSN	Watson	67	185	0.879
N	RBNSN	Robinson	95	267	0.877
N	MRMTK	Marmaduke	27	78	0.873
N	BRYN	Bryan	29	91	0.859
N	WLKNSN	Wilkinson	25	84	0.848
N E	K0BRT	Cuthbert	96	226	0.961
N E	RBSN	Robson	21	58	0.950
N E	TR0	Dorothy	696	2449	0.930
N E	LNSLT	Lancelot	50	180	0.928
N E	BRBR	Barbara	406	1470	0.927
N E	ANBL	Annabella	41	153	0.924
N E	RLF	Ralph	578	2548	0.907
N E	FSTR	Foster	20	91	0.904
N E	0MSN	Thomasina	66	315	0.899
N E	ISBL	Isabella	3190	15396	0.897
N Mid	STKLF	Sutcliffe	67	70	0.986
N Mid	HWR0	Howarth	51	55	0.977
N Mid	KRNWT	Greenwood	122	132	0.973
N Mid	BT	Betty	1816	2118	0.949
N Mid	SKR	Squire	326	409	0.923

N Mid	HRTL	Hartley	160	201	0.923
N Mid	RFT	Wright	263	333	0.921
N Mid	NNS	Nancy	2043	2948	0.874
N Mid	HMLT	Hamlet	57	84	0.869
N Mid	LSTR	Lester	80	119	0.861
N Wales	KRF0	Griffith	480	803	0.973
N Wales	OWN	Owen	889	1869	0.956
N Wales	KWN	Gwen	207	460	0.952
N Wales	HF	Hugh	968	2835	0.926
N Wales	IXML	Ishmael	22	72	0.914
N Wales	HMFR	Humphrey	108	376	0.907
N Wales	EFN	Evan	665	3028	0.872
N Wales	WTKN	Watkin	26	124	0.865
N Wales	LWLN	Llewellyn	63	312	0.860
N Wales	KMR	Gomer	11	59	0.847
S E	TRTN	Trayton	56	69	0.919
S E	FLTLF	Philadelphia	93	126	0.882
S E	LPLT	Leopold	38	56	0.848
S E	WR0	Worthy	39	58	0.844
S E	ERL	Earl	30	51	0.791
S E	AKSTS	Augustus	445	776	0.780
S E	HSTR	Hester	581	1023	0.776
S E	ATLFS	Adolphus	121	214	0.774
S E	BRTRM	Bertram	33	59	0.770
S E	TKLS	Douglas	134	242	0.767
S Wales	KWNLN	Gwenllian	234	247	0.998
S Wales	JNKN	Jenkin	201	216	0.997
S Wales	HWL	Howell	140	170	0.991
S Wales	MRKN	Morgan	529	649	0.990
S Wales	EFN	Evan	2017	3028	0.979
S Wales	KMR	Gomer	38	59	0.977
S Wales	WTKN	Watkin	71	124	0.969
S Wales	ELFR	Elvira	50	97	0.960
S Wales	LWLN	Llewellyn	150	312	0.955
S Wales	KWN	Gwen	214	460	0.953
W	LFT	Loveday	62	96	0.966
W	WLMT	Wilmot	42	101	0.921
W	BS	Bessie	957	2706	0.898
W	NXLS	Nicholas	342	1073	0.883
W	0MSN	Thomasina	100	315	0.883
W	TRFN	Tryphena	68	246	0.861

W	LSNT	Lucinda	28	105	0.856
W	XRT	Charity	91	349	0.850
W	NN	Nina	120	519	0.829
W	KRS	Grace	1077	4728	0.826

	Speaks Welsh		Speaks English	
	(1)	(2)	(3)	(4)
North Wales score	0.355	0.060	−0.178	−0.031
	(0.087)	(0.017)	(0.057)	(0.010)
South Wales score	0.498	0.266	−0.019	−0.020
	(0.048)	(0.035)	(0.022)	(0.005)
Southeast England score	−0.543	−0.246	0.104	0.040
	(0.021)	(0.033)	(0.015)	(0.005)
Birthplace FE		x		x
N	97932	97932	97932	97932
R^2	0.146	0.507	0.023	0.171

This table presents evidence of the relationship between region-specific name scores and speaking Welsh or English in the 1911 census. The language question was only asked in Wales and the Isle of Mann. We exclude those recorded as speaking neither language. Data is aggregated to the name-birthplace district level, observations are weighted by the number of individuals in each cell, to give the same coefficients as an individual-level regression. In (1) and (2) the dependent variable is an indicator that the individual speaks Welsh, in (3) and (4) and indicator that they speak English. The independent variables are the name scores for the individual's name and the North Wales, South Wales and Southeast England cultural regions. Models (2) and (4) add fixed effects for the district of birth, (1) and (3) include and intercept. Standard errors clustered by district of birth in parentheses.

Table B.4: Relationship between name scores and language

Table B.8: Data construction overview

	(1)	(2)	(3)
Task	Allocating districts to clusters	Calculating name scores	Calculating origin-destination-cluster migration flows
Data	Surnames of household heads born before 1800, allocated to district of birth	First names of people born 1841–1860, allocated to cluster of birth	People born 1861–1911, allocated to clusters by first name, to origin districts by birth, to destinations by residence
Source	1851 Census	1911 Census	1911 Census

	distance	religious distance		migrants			marriages			linguistic overlap	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Same cluster	−0.658	−0.185	−0.050	3.430	0.320	0.301	3.063	0.311	0.106	0.414	0.027
	(0.023)	(0.012)	(0.008)	(0.075)	(0.059)	(0.055)	(0.046)	(0.043)	(0.017)	(0.012)	(0.008)
log distance			0.164		−1.660	−1.566		−1.616	−0.387		−0.377
			(0.009)		(0.089)	(0.094)		(0.061)	(0.042)		(0.008)
log religious distance						−0.444			−0.125		
						(0.055)			(0.012)		
log migrants									0.641		
									(0.012)		
Cluster coefficient as %	−48	−17	−5	2986	38	35	2039	36	11	51	3
N	705600	698894	698058	688864	688040	685892	678976	678152	269707	925406	923626
Pseudo R^2	0.439	0.032	0.034	0.417	0.752	0.755	0.433	0.706	0.788	0.045	0.052

This table shows dyadic regressions of measures of similarity between two locations against an indicator that the origin and destination are allocated to the same cultural cluster. Models (1)–(9) are estimated on registration district-dyads. In (1) the dependent variable is the log distance between districts, in (2)–(3) the log Euclidean distance of religious attendance shares in 1851, in (4)–(6) the log number of migrants in the 1851 census, in (7)–(9) the log number of married couples in the 1851 census in which the man is born in the origin and the woman in the destination. Models (10) and (11) are estimated on dyads for geocoded medieval texts in Benskin et al. (2013); the dependent variable is the log share of linguistic features shared between the two texts. All models are estimated by Poisson Pseudo-Maximum Likelihood, with fixed effects for the origin and destination, and standard errors clustered by origin and destination.

Table B.1: Surname clusters correlate with other forms of cultural similarity

	migration		marriage		children's naming	
	(1)	(2)	(3)	(4)	(5)	(6)
Name cluster	0.486	0.138	0.166	0.049	0.579	0.506
	(0.019)	(0.011)	(0.009)	(0.003)	(0.009)	(0.010)
Location-name cluster FE	x	x	x	x	x	x
Cluster of interest FE	x		x		x	
Location-cluster of interest FE		x		x		x
N	66807	65465	53016	53016	65466	65232
Pseudo R^2	0.192	0.484	0.248	0.260	0.252	0.259

This table shows dyadic regressions of how the cultural cluster an individual is assigned bassed on his or her name predicts other cultural behaviors. Data is at the location-individual cluster-cluster of interest level. Models are estimated by Poisson-Pseudo Maximum Likelihood, with fixed effects for each location-name cluster combination and weights for the number of individuals in each location-name cluster group of cells; this gives equivalent estimates to running multinomial logit on individual-level choice data. The independent variable of interest is whether the cluster the individual is allocated to corresponds to the culture of interest. In (1)–(2), the dependent variable is the share of individuals born in the region assigned names associated with a cluster migrating to the cluster of interest, in (3)–(4) the the share of men in the region with names associated with a cluster married to women with names allocated to the cluster of interest, in (5)–(6) the share of children of men in a given district with names associated with a cluster given names allocated to the cluster of interest. In (1)–(2) location is the individual's registration district of birth, in (3)–(6) his district of residence. Odd-numbered models also include fixed effects for the cluster of interest, even-numbered models include fixed effects for location-cluster of interest combinations, these more granular fixed effects adjust for the propensity of people born in a district to migrate to a given cluster (1–2), for the propensity of men resident in a district to marry women given names associated with a given cluster (3–4) and for the propensity of men in a given district to give their children names associated with a given cluster (5–6). Standard errors clustered by location in parentheses.

Table B.3: The cluster an individual is allocated to based on his or her name predicts migration, marriage, and children's naming

	Observed cultural change ($\Delta \ln \sigma$)					
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted cultural change ($\Delta \ln \sigma$)						
from real wages (v_d^θ)	0.739	0.305				
	(0.021)	(0.054)				
from migration costs (δ_{od}^θ)			0.123	0.260		
			(0.036)	(0.027)		
from starting population (n_o)					0.388	0.445
					(0.026)	(0.102)
Cluster FE		x		x		x
N	7362	7362	7362	7362	7362	7362
R^2	0.234	0.521	0.004	0.530	0.043	0.523

This table shows OLS estimates at the district-by-cluster level. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ in (1) and (2), migration costs in (3) and (4), and starting populations in (5) and (6), and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Even-numbered models add cluster fixed effects. Standard errors clustered by district in parentheses.

Table B.5: Relationship between the change in log cultural choice shares, σ, relative to different 1851 counterfactuals

	Change in home culture ($\Delta \ln \sigma$)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln$ real wage (v_d^θ)	−0.032	−0.056	−0.216	−0.161	−0.063	0.003
	(0.027)	(0.026)	(0.050)	(0.035)	(0.023)	(0.009)
Distance to London (100km)		−0.392	−0.867	−0.567	−0.354	−0.112
		(0.102)	(0.123)	(0.073)	(0.065)	(0.052)
$\Delta \ln v_d^\theta \times$ distance		0.078	0.183	0.120	0.081	0.014
		(0.025)	(0.028)	(0.016)	(0.015)	(0.012)
Counterfactual	Realized	Realized	All	Wages	Pop.	Migration costs
N	812	812	812	812	812	812
R^2	0.002	0.031	0.083	0.092	0.027	0.081

This table shows OLS estimates at the district-level of the relationship between wage growth and the popularity of the home culture across counterfactuals. The dependent variable is the log share choosing the home culture minus the log share predicted choosing the home culture in a specific counterfactual scenario. Models (1) and (2) examine the observed change between the 1841–1860 generation and the 1861–1895 generation. (3) examines the output from the counterfactual setting real wages, starting populations and migration costs to 1851 levels, (4) just real wages, (5) just populations, and (6) just migration costs. The independent variable is the change in real wages between 1851 and 1911, interacted with distance to London. All models include an intercept. Robust standard errors in parentheses.

Table B.6: Relationship between wage growth, distance to London, and counterfactual changes in the popularity of the home culture

	Change in log share migrating outside home culture			
	(1)	(2)	(3)	(4)
$\Delta \ln$ real wage (v_d^θ)	−0.125	−0.296	0.035	0.135
	(0.049)	(0.021)	(0.024)	(0.018)
Distance to London (100km)	0.559	0.433	0.011	−0.127
	(0.230)	(0.065)	(0.116)	(0.064)
$\Delta \ln v_d^\theta \times$ distance	−0.100	−0.098	−0.000	0.048
	(0.052)	(0.015)	(0.026)	(0.015)
Counterfactual	All	Wages	Pop.	Migration costs
N	828	828	828	828
R^2	0.145	0.677	0.019	0.317

This table shows OLS estimates at the district-level of the relationship between wage growth and migration from the home culture across counterfactuals. The dependent variable is the log share migrating to locations outside the cluster to which we allocate the district based on historical surnames in the observed data, minus the predicted log share in the relevant counterfactual. (1) examines the output from the counterfactual setting real wages, starting populations and migration costs to 1851 levels, (2) just real wages, (3) just populations, and (64 just migration costs. The independent variable is the change in real wages between 1851 and 1911, interacted with distance to London. All models include an intercept. Robust standard errors in parentheses.

Table B.7: Relationship between wage growth, distance to London, and counterfactual changes in migration from the cluster

Table B.9: Estimation overview

	(1)	(2)	(3)	(4)	(5)
Parameters	δ_{od}^θ, bundled $\left(\left(v_d \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha\right)$	$\alpha, \left(\xi_d^k\right)^\theta, v_d^\theta$	Ω_o^k	φ, ψ_o^k	$v_{d,1851}^\theta, \delta_{od,1851}^\theta$
Estimation	Poisson regression	Two-stage least squares	Calculated directly from (1) and (2)	Poisson regression with control function	Poisson regression
Level	Origin-destination-culture	Destination-culture	Origin-culture	Origin-culture	Origin-destination
Data	1911 census	1911 census	1911 census	1911 census	1851 census

C EQUILIBRIUM UNIQUENESS

Allen, Arkolakis and Li (2020) provide conditions for equilibrium existence and uniqueness in spatial economic models. Their analysis focuses on an economic model with N agents and H types of interactions in which the equilibrium can be reduced to a set of $N \times H$ equations of the form

$$x_{ih} = \sum_{j=1}^{N} f_{ijh}(x_{j1}, ..., x_{jH}) \tag{6}$$

where x_{ih} is the equilibrium outcome for each agent in each interaction and f_{ijh} is a differentiable function that governs the interactions between agents. One can construct an $H \times H$ matrix of uniform bounds of the elasticities of these f_{ijh} functions,

$$\mathbf{A}_{hh'} = \sup_{i,j} \left(\left| \frac{\partial \ln f_{ijh}}{\partial \ln x_{jh'}} \right| \right)$$

Theorem 1 of Allen, Arkolakis and Li (2020) states that if the spectral radius of this matrix—the largest absolute value of its eigenvalues—is less than one, the equilibrium exists, is unique, and can be computed by iterating equation (6).

The main theorem requires that the f_{ijh} function only takes as arguments equilibrium outcomes pertaining to agent or location j and returns strictly positive values. Remark 1 gives an additional result for cases where f_{ijh} takes as arguments equilibrium outcomes pertaining to other agents or locations as well, and returns zeros. The results of Theorem 1 hold in such cases if one replaces the uniform bound on the elasticity with the uniform bound on the sum of elasticities.

To apply this theorem, we need to rewrite the equilibrium conditions of our model in a form consistent with equation (6).

Note that we can write (1), the number with culture k migrating from o to d, as

$$m_{od}^k = \frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\Omega_o^k} \frac{\left(\Omega_o^k\right)^\varphi \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l} n_o = \left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha n_o \frac{\left(\Omega_o^k\right)^{\varphi-1} \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l}$$

we can therefore write the total number of migrants of culture k in d as

$$m_d^k = \sum_{o=1}^{N} \left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha n_o \frac{\left(\Omega_o^k\right)^{\varphi-1} \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l}$$

Defining $q_d^k = \left(m_d^k\right)^{1-\alpha}$, we can rewrite this equation as

$$q_d^k = \sum_{o=1}^{N} \left(v_d \delta_{od} \xi_d^k\right)^\theta n_o \frac{\left(\Omega_o^k\right)^{\varphi-1} \psi_o^k}{\sum_{l=1}^K \left(\Omega_o^l\right)^\varphi \psi_o^l}$$

Similarly we can rewrite

$$\Omega_o^k = \sum_{j=1}^{N} \left(v_j \delta_{oj} \xi_j^k\right)^{\theta} \left(q_j^k\right)^{\frac{\alpha}{1-\alpha}}$$

We can thus express our equilibrium in terms of a set of $N \times K \times 2$ equations of the form $\Omega_o^k = \sum_{j=1}^{N} \sum_{l=1}^{K} g_{oj}^{k,l}(\mathbf{q}, \mathbf{\Omega})$ and $q_o^k = \sum_{j=1}^{N} \sum_{l=1}^{K} h_{oj}^{k,l}(\mathbf{q}, \mathbf{\Omega})$, where

$$g_{oj}^{k,l} = \left(v_j \delta_{oj} \xi_j^l\right)^{\theta} \left(q_j^l\right)^{\frac{\alpha}{1-\alpha}} \mathbf{1}_{\{l=k\}},\ h_{oj}^{k,l} = \left(v_o \delta_{jo} \xi_o^l\right)^{\theta} n_j \frac{\left(\Omega_j^l\right)^{\varphi-1} \psi_j^l}{\sum_{m=1}^{K} \left(\Omega_j^m\right)^{\varphi} \psi_j^l} \mathbf{1}_{\{l=k\}}.$$

In relation to Allen, Arkolakis and Li (2020)'s notation, here we treat $N \times K$ location-by-culture pairs as the set of locations and Ω_o^k and q_o^k as the two types of equilibrium outcomes. Note that the $g_{oj}^{k,l}$ function depends only on q_l^k, but can return zero values if $k \neq l$, and so is not consistent with equation (1) of Allen, Arkolakis and Li (2020). The $h_{oj}^{k,l}$ function depends on both Ω_j^l and Ω_j^m for other values $m \neq l$. We must apply Remark 1 of Allen, Arkolakis and Li (2020) which gives a condition related not to the uniform bound on the elasticity but to the uniform bound on the sum of the elasticities.

The relevant matrix of uniform bounds on the elasticities is then

$$\mathbf{A} = \begin{pmatrix} \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m}}{\partial \ln q_j^l} \right| & \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m}}{\partial \ln \Omega_j^l} \right| \\ \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} g_{od}^{k,m}}{\partial \ln q_j^l} \right| & \sup_{o,k} \sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln \sum_{m=1}^{K} \sum_{d=1}^{N} g_{od}^{k,m}}{\partial \ln \Omega_j^l} \right| \end{pmatrix}$$

For the top right-hand corner, note that $\sum_{m=1}^{K} \sum_{d=1}^{N} h_{od}^{k,m} = \sum_{d=1}^{N} h_{od}^{k,k} = q_o^k$

$$\frac{\partial \ln q_o^k}{\partial \ln \Omega_j^l} = \frac{\partial \ln q_o^k}{\partial \ln h_{oj}^{k,k}} \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} = \frac{h_{oj}^{k,k}}{q_o^k} \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l}$$

Note,

$$\frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^k} = \varphi - 1 - \sigma_j^k \varphi,\ \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} = -\sigma_j^l \varphi$$

where σ_j^l is the share choosing culture l in location j. The sum of these absolute values is then

$$\sum_{j=1}^{N} \sum_{l=1}^{K} \left| \frac{\partial \ln q_o^k}{\partial \ln \Omega_j^l} \right| = \sum_{j=1}^{N} \frac{h_{oj}^{k,k}}{q_o^k} \left(\sum_{l=1}^{K} \left| \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} \right| \right)$$

Writing out the term in parentheses

$$\sum_{l=1}^{K} \left| \frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l} \right| = \left| \varphi - 1 - \sigma_j^k \varphi \right| + \sum_{l \in \{1,\ldots,K \setminus k\}} \sigma_j^l \varphi$$

Given that $\sum_{l\in\{1,\ldots,K\setminus k\}}\sigma_j^l = 1-\sigma_j^k$, we can rewrite this equation as

$$\sum_{l=1}^{K}\left|\frac{\partial \ln h_{oj}^{k,k}}{\partial \ln \Omega_j^l}\right| = \left|\varphi\left(1-\sigma_j^k\right)-1\right| + \varphi\left(1-\sigma_j^k\right)$$

If $\varphi\left(1-\sigma_j^k\right)\leq 1$, this expression equals 1. If $\varphi\left(1-\sigma_j^k\right)>1$, it equals $2\varphi\left(1-\sigma_j^k\right)-1$, which is maximized at $\sigma_j^k=0$. This expression is therefore bounded by $\max\left(2\varphi-1,1\right)$ So we have

$$\sum_{j=1}^{N}\sum_{l=1}^{K}\left|\frac{\partial \ln q_o^k}{\partial \ln \Omega_j^l}\right| \leq \sum_{j=1}^{N}\frac{h_{oj}^{k,k}}{q_o^k}\max\left(2\varphi-1,1\right) = \max\left(2\varphi-1,1\right)$$

For bottom-left entry in $\mathbf{A}$,

$$\sum_{m=1}^{K}\sum_{d=1}^{N} g_{od}^{k,m} = \sum_{d=1}^{N} g_{od}^{k,k} = \Omega_o^k$$

$$\frac{\partial \ln \Omega_o^k}{\partial \ln q_j^k} = \frac{\partial \ln \Omega_o^k}{\partial \ln g_{od}^{k,k}}\frac{\partial \ln g_{od}^{k,k}}{\partial \ln q_j^k} = \frac{g_{od}^{k,k}}{\Omega_o^k}\frac{\alpha}{1-\alpha}$$

Note that $\frac{\partial \ln \Omega_o^k}{\partial \ln q_j^l}=0$ for $l\neq k$. Then

$$\sum_{j=1}^{N}\sum_{l=1}^{K}\left|\frac{\partial \ln \Omega_o^k}{\partial \ln q_j^l}\right| = \sum_{j=1}^{N}\left|\frac{\partial \ln \Omega_o^k}{\partial \ln q_j^k}\right| = \sum_{j=1}^{N}\frac{g_{od}^{k,k}}{\Omega_o^k}\frac{\alpha}{1-\alpha} = \frac{\alpha}{1-\alpha}$$

The two diagonals are zero because $\frac{\partial h_{od}^{k,m}}{\partial q_j^l}=0$ and $\frac{\partial g_{od}^{k,m}}{\partial \Omega_j^l}=0$ for all o,d,k,m,j,l.

We can therefore write

$$\mathbf{A} = \begin{pmatrix} 0 & \max(1,2\varphi-1) \\ \left|\frac{\alpha}{1-\alpha}\right| & 0 \end{pmatrix}$$

The spectral radius of this matrix is $\rho(\mathbf{A}) = \sqrt{\left|\frac{\alpha}{1-\alpha}\right|\max(1,2\varphi-1)}$. The condition that $\rho(\mathbf{A})<1$ will be satisfied if $\frac{\alpha}{1-\alpha}\max(1,2\varphi-1)<1$

D ADDITIONAL DETAIL ON INSTRUMENTAL VARIABLES

D.1 φ *Equation*

We construct the instrument as follows. First, we aggregate migration data at the origin-destination level, ignoring clusters, and regress log migration flows against origin and destination fixed effects, log distance, and an indicator for the origin being the same as the destination:

$$\ln m_{od} = \gamma_d + \beta_1 \ln \text{distance}_{od} + \beta_2 \mathbf{1}_{\{o=d\}} + \gamma_o + \varepsilon_{od}.$$

This gravity regression gives an alternative set of migration cost coefficients, β_1, β_2, and destination fixed effects γ_d. We generate predicted destination fixed effects by regressing the estimated destination fixed effects against an indicator for coal presence:

$$\gamma_d = \delta_0 + \delta_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \varepsilon_d.$$

These estimates allow us to predict the migration pull to different locations, using just the presence of coal and geographic distance. We use these predicted migration incentives to construct on instrument for Ω_o^k:

$$\text{coal-predicted } \Omega_o^k = \sum_{d=1}^{N} \exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{od} + \hat{\beta}_2 \mathbf{1}_{\{o=d\}}\right) \mathbf{1}_{\{\text{historic culture}(d)=k\}}$$

where $\hat{v}_d^\theta$ is the exponential of the coal-predicted destination utility, and the indicator function on the right takes a value of 1 if the district d is allocated to cluster k based on surnames before 1800. The logic is that coal and distance provide pull factors unrelated to cultural sorting, while locations allocated to a specific culture due to historic surnames will have higher cultural amenities ξ_d^k for that culture.

Again following Borusyak and Hull (2020) we recenter the instrument by permuting the vector of coal deposits 1,000 times, calculating Ω_o^k on each permutation using the permuted coal deposits, and subtracting the average $\ln \Omega_o^k$ over all permutations from our instrument. Recentering helps address the concern that districts close in space to a given historic culture may also have more similar cultural transmission tastes.

We then estimate φ with the control function as follows:

$$\ln \Omega_o^k = \beta \text{ recentered} \ln \text{coal-predicted } \Omega_o^k + \gamma_o + \gamma_{k(o)}^k + e_o^k$$

$$\ln \sigma_o^k = \varphi \ln \Omega_o^k + \delta\left(\ln \Omega_o^k - \ln \hat{\Omega}_o^k\right) + \gamma_o + \gamma_{k(o)}^k + \ln \psi_o^k$$

where $\gamma_{k(o)}^k$ is a fixed effect for the surname culture cluster o is located in interacted with the culture in question k, and $\left(\ln \Omega_o^k - \ln \hat{\Omega}_o^k\right)$ is the first-stage residual. The $\gamma_{k(o)}^k$ fixed effects help further address the concern that districts close to a given surname cluster might have stronger cultural transmission tastes for that cluster. We use our estimate of φ and observed σ_d^k and Ω_o^k to back out ψ_o^k.

D.2 α *Equation*

The coal-related influence on cultural choices that we use to instrument for Ω_o^k can also be used to instrument for destination populations and estimate α. We can predict migration flows from origins to destinations using coal, distance, and starting populations:

$$\text{coal-predicted } m_{od} = \frac{\exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{d \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{od} + \hat{\beta}_2 \mathbf{1}_{\{o=d\}}\right)}{\sum_{j=1}^{N} \exp\left(\hat{\delta}_0 + \hat{\delta}_1 \mathbf{1}_{\{j \text{ contains coal}\}} + \hat{\beta}_1 \ln \text{distance}_{oj} + \hat{\beta}_2 \mathbf{1}_{\{o=j\}}\right)} n_o.$$

We then use these predicted migration flows—which do not take into account cultural differences—to predict how shocks to choices in the origins translate into destination populations:

$$\text{coal-predicted } m_d^k = \sum_{o=1}^{N} \text{coal-predicted } \Omega_o^k \cdot \text{coal-predicted } m_{od}.$$

This measure gives the weighted sum of shocks to cultural choices at the origin, weighted by the predicted number of migrants from that origin.

We also recenter this instrument following Borusyak and Hull (2020). As for the instrument for Ω_o^k, we permute the vector of coal deposits 1,000 times, calculate the instrument on each permutation, and substract the average of $\ln$ coal-predicted m_d^k from our instrument. When recentering the instrument, we only permute the coal deposits used to calculate coal-predicted Ω_o^k, not those used to calculate migration flows. Thus the average of the permuted instruments takes into account how coal infuences migration to specific origins, and so subtracting that average nets out such an effect. We then estimate the following equations by two-stage least squares:

$$\ln m_d^k = \beta \text{ recentered} \ln \text{coal-predicted } m_d^k + \gamma_d + \gamma_{k(d)}^k + e_d^k$$

$$\gamma_d^k = \alpha \ln \hat{m}_d^k + \gamma_d + \gamma_{k(d)}^k + \theta \ln \xi_d^k.$$

D.3 *Validating Instruments*

The exclusion restrictions necessary for our instrumental variables estimations are that the instrument for destination populations is uncorrelated with cultural amenities (ξ_d^k), and that the instrument for origin expected utilities is uncorrelated with cultural transmission tastes (ψ_o^k). While we cannot directly test that assumption, we can examine the relationship between our instruments and measures of cultural similarity that should be correlated with the taste and amenity terms. We examine the religious distance and surname similarity measures used in Section 6.1, and the average distance to districts allocated to the cluster in question, which should pick up other forms of cultural similarity that are geographically clustered.

Table D.1 regresses these measures of cultural similarity against our instruments, as in the first stages of our instrumental variables estimation routines. In all models, the coefficient on the instrument is close to zero and not statistically significant, which raises our confidence in the exclusion restriction.

	ln religious distance		ln surname similarity		ln average distance	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	−0.015		0.734		−0.200	
	(0.056)		(0.584)		(0.180)	
Recentered coal-predicted $\ln \Omega_o^k$		0.023		0.273		−0.115
		(0.044)		(0.418)		(0.107)
N	7425	7425	7264	7264	7452	7452
R^2	0.842	0.848	0.906	0.907	0.845	0.859

This table validates the exclusion restriction assumption for our instruments used to estimate α and φ by examining the relationship between the instruments and religious distance and surname similarity, measures of cultural proximity that correlate with ξ and ψ, the residuals in the second stage regressions, as well as average geographic distance to districts allocated to the culture in question. Data is at the district-cluster level. As in Table 3, observations are weighted by the relevant origin or destination population, and all models include district and cluster-by-home-cluster fixed effects. Standard errors clustered by district in parentheses.

Table D.1: Orthogonality of Instruments to Cultural Proximity

E BOOTSTRAPPING ELASTICITIES

Table E.1 reports bootstrapped confidence intervals for the homophily (α) and culture (φ) elasticities. In our model, there are two sources of uncertainty. First, there uncertainty from the sampling of locations into our dataset. Second, origin-destination-culture migration flows are affected by transitory shocks that we treat as an error term in the estimation of distance elasticities and destination bundles of attributes. One can think of migration flows as a draw from a multinomial distribution. Randomness in the realization of this draw will affect our estimates.

To account for sampling uncertainty, one would normally use the nonparametric bootstrap, blocking by location. That is, one would resample locations or weights for locations to obtain a set of new datasets, and then run the entire estimation routine on each new dataset to obtain the sampling distribution of the quantity of interest. The key idea is that each resampled dataset approximates a draw from the same data generating process as the primary dataset. In our context that assumption would not be satisfied. Our model and counterfactual exercises indicate that changing the set of locations or populations alters individuals' cultural choices and migration decisions. If a bootstrap replication sampled London twice, we would expect more people to migrate there and the popularity of the southeast to increase. Observed cultural choice and migration shares that do not take into account this response would be inconsistent with the model.

Our preferred approach combines the clustered sampling aspect of the nonparametric bootstrap with the structure of our model. On each bootstrap replication we do the following:

1. Sample weights for origin locations from a Dirichlet distribution, as in a fractional random weight bootstrap

	$\ln m_d^k$	γ_d^k		$\ln \Omega_o^k$	$\ln \sigma_o^k$	
	(1)	(2)	(3)	(4)	(5)	(6)
Recentered coal-predicted $\ln m_d^k$	0.641					
	[0.095; 2.564]					
$\ln m_d^k$ [α]		0.604	0.521			
		[0.571; 0.634]	[0.385; 0.873]			
Recentered coal-predicted $\ln \Omega_o^k$				0.214		
				[0.070; 0.718]		
$\ln \Omega_o^k$ [φ]					2.185	1.803
					[2.038; 2.390]	[0.844; 2.355]
First-stage residuals						0.388
						[−0.040; 3.510]
Model	First stage	OLS	TSLS	First stage	PPML	Control function
N	7418	7418	7418	7452	7452	7452
R^2	0.989	0.999	0.999	0.998		

This table shows estimates of α and φ with confidence intervals from a semiparametric bootstrap. All models are estimated at the district-by-culture level. Models (1) and (4) show the first stage, (2) and (5) the non-instrumented second stage and (3) and (6) the instrumented second stage. (1), (2), and (4) are estimated by OLS, (3) by TSLS, and (5) and (6) by Poisson Pseudo-Maximum Likelihood, with (6) implementing the control function approach by including the residuals from (4). In models (2) and (3) the dependent variable is the destination-by-culture fixed effect from an origin-by-destination-by-culture gravity model, in (5) and (6) it is the log share of people born in the district assigned to the culture. All models include fixed effects for the district and the culture interacted with the historic culture of the district. Standard errors are calculated through a semiparametric bootstrap clustered by location. On each bootstrap iteration we sample weights for origin locations, solve for equilibrium migration and cultural choices, draw origin-destination-culture migration flows from a multinomial distribution given these choices and run the entire estimation routine on the resulting dataset. Confidence intervals are the 2.5th and 97.5th percentiles of the bootstrap distribution.

Table E.1: Estimates of model elasticities with bootstrapped confidence intervals

2. Multiply origin populations by these sampled weights
3. Solve for equilibrium migration flows and cultural choices given these new origin populations
4. Sample migration flows from a multinomial distribution, using the predicted share of members of culture k going from o to d as the choice probabilities
5. Estimate a gravity model on these sampled migration flows to estimate destination fixed effects and migration cost parameters
6. Calculate instruments using these sampled migration flows
7. Use calculated instruments to estimate elasticities.

Doing so 1,000 times gives a distribution of estimated elasticities over resamples. We construct confidence intervals using the 2.5th and 97.5th percentiles of these distributions in Table E.1.

F ESTIMATING 1851 ECONOMIC FUNDAMENTALS

This appendix details how we estimate real wages and migration costs using 1851 data. In the main estimation, we use origin-destination-culture migration flows to separately estimate the real wage v_d^θ and cultural amenities ξ_d^k. For 1851, we do not have data on names and so cannot allocate individuals to cultures. We estimate the following gravity model using origin-destination migration flows from 1851:

$$\ln m_{od,1851} = \gamma_{d,1851} + \beta_{1,1851} \ln \text{distance}_{od} + \beta_{2,1851} \mathbf{1}_{\{o=d\}} + \gamma_{o,1851} + \varepsilon_{od,1851} \tag{7}$$

The parameters $\beta_{1,1851}$ and $\beta_{2,1851}$ pin down migration costs as a function of geographic distance and an indicator that the origin is the same as the destination, much as in our main estimation in Section 5.1.

In a model without cultural differentiation, the destination fixed effect $\gamma_{d,1851}$ would correspond to $\theta \ln v_{d,1851}$, the real wage. However, in our model agents have a preference for being in a location with other members of the same culture, so some of the migration pull to a particular location is due to the number of others migrating there. To estimate the real wage, we need to rescale these destination fixed effects. We estimate the equivalent of (7) using 1911 data, and then regress the log real wages, $\theta \ln v_d$, estimated for 1911 in Section 5.2 against $\gamma_{d,1911}$, the destination fixed effects from this 1911 gravity model. In that regression, the coefficient on the destination fixed effects is 0.43 and the R^2 is 0.88. We then rescale the destination fixed effects from the 1851 model by 0.43 to account for this sorting effect: $v_{d,1851}^\theta = \exp(\gamma_{d,1851})^{0.43}$

G WHY THE SOUTHEAST?

Why was it specifically the culture of the Southeast that grew over the 1851–1911 period? Table 5 shows that going from 1911 economic parameters to 1851 parameters leads to a 28% decrease in the size of the Southeast English culture, after accounting for the mechanical effects of changing population origins. This change is almost entirely due to the effects of changing real wages across locations. In this appendix, we use the model to illustrate how the relationship between real wage changes, cultural amenities, and population geography account for this pattern.

In the model, the number of people choosing a culture k is the sum over locations of the shares of people choosing the culture in that location (σ_o^k) multiplied by the starting population (n_o):

$$\textit{Culture size}^k = \sum_{o=1}^{N} \sigma_o^k n_o = \sum_{o=1}^{N} n_o \frac{\left(\Omega_o^k\right)^\varphi \psi_o^k}{\sum_{l=1}^{K} \left(\Omega_o^l\right)^\varphi \psi_o^l}.$$

Changes to real wages can influence the overall sizes of cultural groups by changing Ω_o^k across locations, especially in places with large origin populations and by increasing Ω_o^k relative to the values of Ω_o^l for other cultures $l \neq k$.

Ω_o^k is a sum of real wages multiplied by cultural amenities, destination populations, and

migration costs:

$$\Omega_o^k = \sum_{j=1}^{N} \delta_{oj}^{\theta} \underbrace{\left(v_j \xi_j^k\right)^{\theta} \left(m_j^k\right)^{\alpha}}_{\text{destination-level terms}} .$$

Note that the $\left(v_j \xi_j^k\right)^{\theta} \left(m_j^k\right)^{\alpha}$ terms vary only at the destination level, while δ_{oj}^{θ} varies by origin and destination.

For intuition about how wage growth across regions can reinforce or undermine cultures, we calculate the partial equilibrium elasticities of culture size to wage growth in each location. These hold fixed the general equilibrium components of the model, whereby changes to wages influence destination populations which in turn influence migration and cultural decisions.

The elasticity of total culture size to wages in a given location can be written as the sum of elasticities of each location's cultural choices to wages in that location:

$$\textit{culture-wage elasticity}_d^k := \frac{\partial \ln \sum_{o=1}^{N} n_o \sigma_o^k}{\partial \ln v_d^{\theta}} = \frac{\sum_{o=1}^{N} n_o \sigma_o^k \frac{\partial \ln \sigma_o^k}{\partial \ln v_d^{\theta}}}{\sum_{o=1}^{N} n_o \sigma_o^k}.$$

In turn, each location-specific choice elasticity can be written as the elasticity of Ω_o^k for the culture in question to wages in a given location, relative to the elasticities of Ω_o^l for other cultures l:

$$\frac{\partial \ln \sigma_o^k}{\partial \ln v_d^{\theta}} = \varphi \frac{\partial \ln \Omega_o^k}{\partial \ln v_d^{\theta}} - \varphi \sum_{l=1}^{K} \sigma_o^l \frac{\partial \ln \Omega_o^l}{\partial \ln v_d^{\theta}}$$

Note that $\frac{\partial \ln \Omega_o^k}{\partial \ln v_d^{\theta}} = \frac{\delta_{od}^{\theta} \left(v_d \xi_d^k\right)^{\theta} \left(m_d^k\right)^{\alpha}}{\Omega_o^k}$. The partial equilibrium elasticity of Ω_o^k to v_d^{θ} is equivalent to the share of members of k originating in o who migrate to d. Inserting this identity gives:

$$\frac{\partial \ln \sigma_o^k}{\partial \ln v_d^{\theta}} = \varphi \delta_{od}^{\theta} v_d^{\theta} \left(\frac{\left(\xi_d^k\right)^{\theta} \left(m_d^k\right)^{\alpha}}{\Omega_o^k} - \sum_{l=1}^{K} \sigma_o^l \frac{\left(\xi_d^l\right)^{\theta} \left(m_d^l\right)^{\alpha}}{\Omega_o^l} \right).$$

Wage growth in d increases k in o if d attracts a large share of migrants of culture k from o, but not a large share of migrants of other cultures in o.

Given these identities we can calculate the partial equilibrium culture-wage elasticity for each culture-location combination. Figure G.1 shows this for all cultures. The general pattern is intuitive. Wage growth in the home region (e.g. North Wales for the N Wales culture) serves to increase the prevalence of the culture nationally, because places in the home region have high amenities and destination populations for the culture, and because they are important migration destinations for members of that culture but not for other cultures.

Returning to the question of why the Southeast grows during the period we study, this exercise shows that it really depends on which regions grow. Our model suggests that the Southeast stands to gain from growth especially around London, relative especially to Lancashire (N Mid), South Wales, and the North East. Comparing the Southeast ("S E") and North Mid ("N Mid") maps, we see that wage growth around London increases the size of the Southeast culture and decreases the North Mid culture, while wage growth in Lancashire (the blue dot in the N Mid map) has the opposite effect. This pattern reflects members of the

Southeast culture originating in various locations tending to migrate to London and members of the North mid culture tending to migrate to Lancashire: wage growth in London raises the relative appeal of migrating to London and thus of belonging to the Southeast culture, relative to the North mid culture.

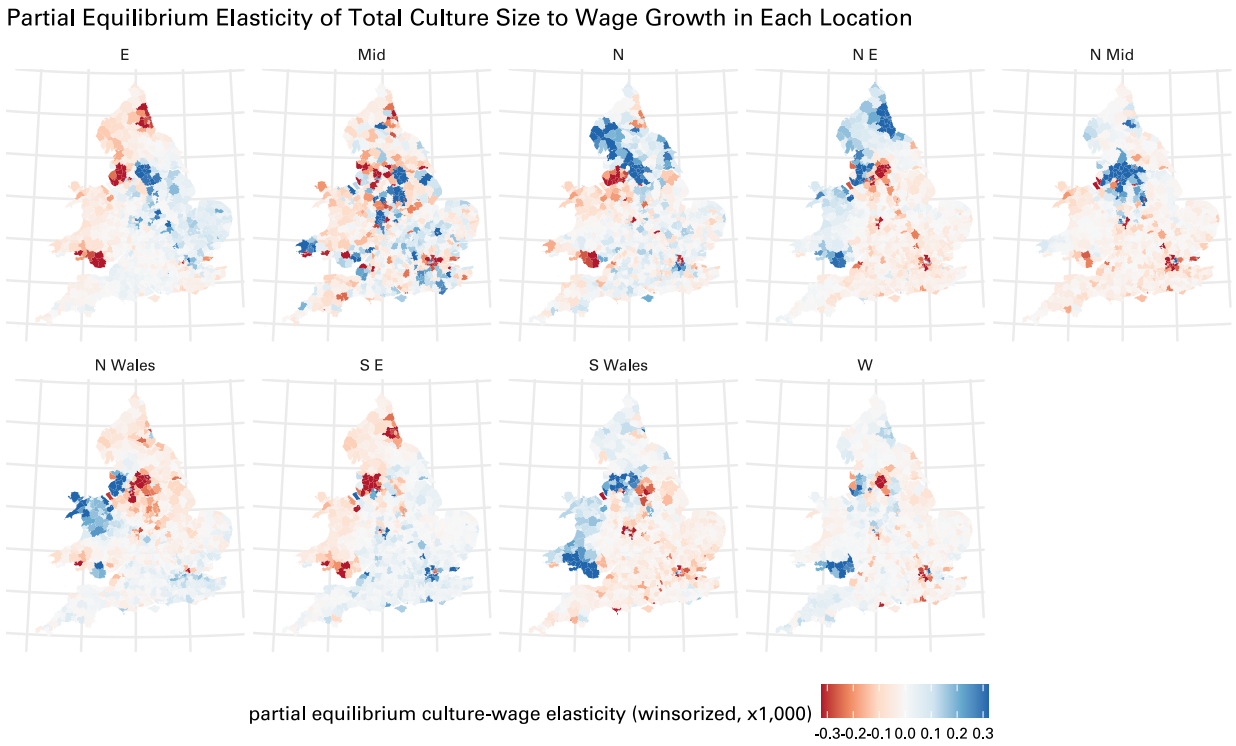


Figure G.1: How much wage growth in each region would increase the log population of each culture

Each map shows the model's partial equilibrium prediction of how wage growth in each region would change the log total population of the culture in question.

Figure G.2 plots the difference between log culture size in the observed data and in the counterfactual using 1851 wages against the correlation between wage growth and culture-wage elasticities. The Southeast culture has the strongest correlation between wage growth over this period and culture-wage elasticities. This correlation largely reflects wage growth around London and slower growth in regions like Lancashire that had been heartlands of the First Industrial Revolution. Areas like the Northeast and South Wales coalfields that industrialized during this period should push against the growth of the Southeast, but Figure G.2 suggests that effect was more than offset by the growth of the area around London. Regions like South Wales and the Northern Midlands that declined during this period have a negative correlation between wage growth and culture-wage elasticities.

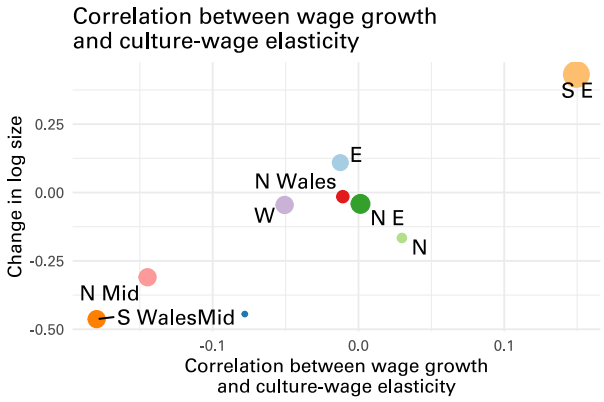


Figure G.2: Cultures for which wage growth correlates with population wage elasticities are predicted greater growth

The x axis show the district-level correlation between real wage growth 1851–1911 and the partial equilibrium elasticity of the culture in question's population to wage growth. A higher value indicates wage growth tended to occur in place where the model would predict that wage growth would reinforce the culture in question. The y axis shows the log change in culture population between the observed data and the counterfactual setting wages to 1851 levels. The size of the dots corresponds to the number allocated to the culture in the observed data.

H COUNTERFACTUAL EFFECTS OF POPULATION CHANGE

Section 6.2 presents the results of counterfactual simulations that set real wages and migration costs to 1851 levels. This appendix present the results of an additional counterfactual that sets origin populations—the third type of economic fundamental in the model—to 1851 levels.

Changing origin populations complement the effects of changing real wages on cultural choices. The third row of Table 5 indicates that going from 1911 to 1851 origin populations decreases the share choosing the Southeast culture by 15%, after netting out the mechanical effect on total populations from allocating population to places where people are more or less likely to choose a given culture. The logic for this effect is that population growth between 1851 and 1911 was larger in areas with higher cultural transmission tastes for the Southeast. More parents choosing to assign their children the Southeast culture in turn creates more migrants belonging to the Southeast culture and larger destination populations for the Southeast culture, which feed into the cultural choices of other parents.

As with wage changes, changing origin populations increased the prevalence of the Southeast culture everywhere (not specifically on the peripheries, see Figure H.1, left panel), and helped preserve coal-producing home cultures (Figure H.1, right panel). Population growth was highest in coal-producing districts and those close to London (Figure A.8, right panel). By influencing destination populations and thus others' migration decisions, population growth served to decrease out-migration from these kinds of areas (Figure A.11D).

Table H.1: Counterfactual estimates

	Home culture pop.		S E Culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	3.193	-0.611	-31.448	-28.323	-35.670	-30.983
1851 real wages, v_d^θ	-16.692	-16.692	-35.070	-35.070	-1.562	-1.562
1851 starting populations	-9.272	-7.048	-22.382	-14.625	-0.782	0.561
1851 migration costs	16.117	16.117	-1.058	-1.058	-31.286	-31.286

This table shows percentage changes in the population choosing each district's home culture, the population choosing the Southeast English culture, and the share migrating from their district of birth. In each group, the first column gives the total percentage change in the quantity of interest under the counterfactual relative to the observed outcome, the second gives the percentage change implied by the share in each location choosing it, fixing origin populations at the observed level. To fix populations, we calculate the weighted average of the quantity of interest in the counterfactual over locations weighted by realized starting population, and dividing that by the weighted average of the realized quantity of interest, weighted by realized starting population. The first row replaces destination real wages, migration costs and starting populations with their estimated 1851 values, the second replaces only destination real wages, the third replaces starting populations, and the fourth migration costs with those estimated from 1851 data.

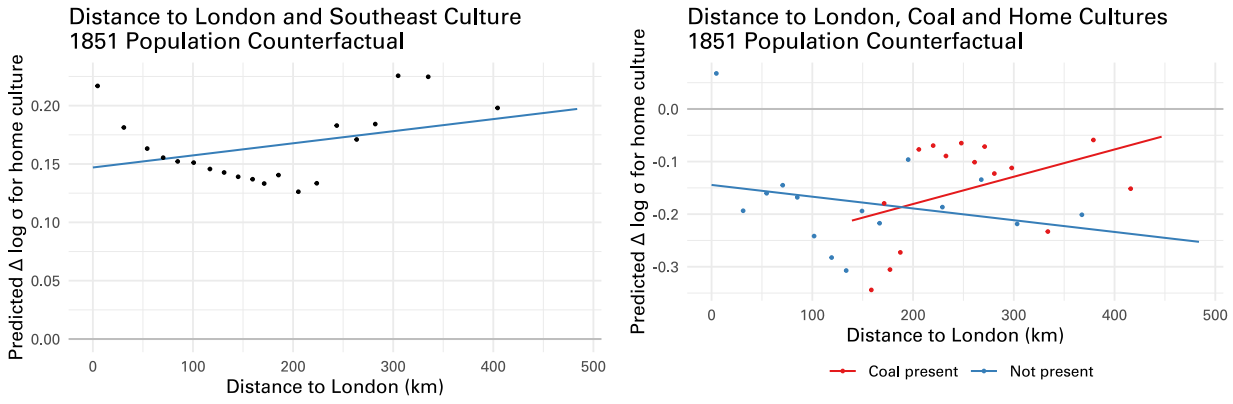


Figure H.1: Counterfactual effects of changing origin populations to 1851 levels

The left panel plots the observed Southeast culture log share minus the log share predicted by a counterfactual setting origin populations to 1851 levels, against distance to the City of London. The right panel plots the change in the log share choosing the home culture in the same counterfactual, split by whether the district contains coal.

I MODEL IMPLIED ELASTICITIES

Does industrialization support or undermine local cultures? Our analysis in the main body of the paper suggests that economic growth has uneven effects across geography, shoring up home cultures in the periphery, but undermining them in the center. In this appendix we use the model to examine how shocks to each locations lead to changes in the popularity of different cultures. This approach differs from Section 6.2, which examined the effects of the specific bundle of realized changes to all locations. For each location we simulate a 1% increase to wages in that location, record the changes to cultural choices in that location, and then convert these changes into elasticities by taking the logarithm and dividing by ln(1.01). These elasticities are the model's prediction of how local economic growth should change the popularities of different cultures in a location. The appendix first validates these elasticities, then discusses how they exhibit a core-periphery pattern, reinforcing the conclusion that economic growth reinforces local cultures on the periphery. Finally we decompose the elasticities to explain why we see this pattern.

Table I.1 examines whether model-implied elasticities moderate the observed relationship between changes in real wages and cultural choices. These elasticities should tell us whether and how much a culture grows in a district in response to economic change in the district. We test this claim by regressing the change in log cultural choices for each culture and district against the change in log real wages in that district, which we interact with the model-implied elasticity. The coefficient of interest is on the interaction between changes in real wages and the elasticity. A positive coefficient would indicate that increasing real wages leads to increases in the prevalence of the cultures with positive culture-wage elasticities. In models (1)–(3) we find this to be the case. In (5) we find that this effect also holds when subsetting to the home culture, the culture to which we allocate the district based on historical surnames. In (4) we find that wage growth in districts further from London preserved the home culture, in (6) we find that this relationship is driven by greater home-culture wage elasticities further from London.

How cultural choices respond to economic change depends upon geography. As shown in Figure I.1, in locations on the periphery, an increase in v_d^θ bolsters the local culture and reduces the appeal of the Southeast culture. In the center, local industrialization shores up the culture of the Southeast, in some cases at the expense of the home culture. Figure I.2 plots these elasticities against distance to London and provides more direct evidence of this core-periphery pattern. As one goes further from the center, shocks to real wages increase the home culture and decrease the Southeast culture. In Table I.1, we find that the tendency of wage growth in districts further from London to preserve the home culture is attributable to the greater model-implied elasticity for the home culture further from London.

What explains the variation in how cultural choices respond to local economic change? While we cannot derive closed-form expressions for the general equilibrium elasticities, we can analyze the elasticities predicted by the model holding some of the general equilibrium components fixed. Log-differentiating equation (3) gives the following expression for the elasticity of σ_o^k to v_o^θ:

$$\frac{\partial \ln \sigma_o^k}{\partial \ln v_o^\theta} = \varphi \frac{\partial \ln \Omega_o^k}{\partial \ln v_o^\theta} - \varphi \sum_{l=1}^{K} \sigma_o^l \frac{\partial \ln \Omega_o^l}{\partial \ln v_o^\theta}.$$

	Change in cultural choice $(\Delta \ln \sigma) \times 100$					
	(1)	(2)	(3)	(4)	(5)	(6)
Model-implied elasticity	−2.122	−1.740	−1.533		−8.044	−8.022
	(0.813)	(0.743)	(0.635)		(2.581)	(2.654)
Distance to London (km)				−0.392		−0.236
				(0.102)		(0.098)
Δ ln real wage $(\Delta\, \theta \ln v_d)$	2.091	1.968		−5.574	−3.830	−2.572
	(0.692)	(0.698)		(2.575)	(2.343)	(2.065)
$\Delta\, \theta \ln v_d \times$ elasticity	0.427	0.330	0.309		1.867	1.912
	(0.189)	(0.176)	(0.151)		(0.639)	(0.656)
$\Delta\, \theta \ln v_d \times$ distance to London				0.078		0.039
				(0.025)		(0.023)
Clusters	All	All	All	Home	Home	Home
Cluster FE		x	x			
District FE			x			
N	7362	7362	7362	812	812	812
R^2	0.006	0.522	0.586	0.031	0.029	0.058

This table shows OLS estimates at the district-by-cluster level. The dependent variable is the change in the log share choosing the culture in the district, between the 1841–1860 cohort and the 1861–1895 cohort, multiplied by 100 to aid legibility. Models (1)–(3) are estimated using all clusters and interact changes in real wages with the model-implied elasticity of σ to v_d^θ, estimated by simulating a 1% increase in v_d^θ separately for each district. The coefficient of interest is this interaction, which should be positive. Models (2) and (3) include cluster fixed effects, (3) includes district fixed effects which are collinear with the change in real wages. Models (4)–(6) subset to the culture cluster the district is allocated based on historical surnames. (4) shows that further from London, wage growth was more positively associated with the preservation of the home culture, (5) shows the interaction between wage growth and the model-implied elasticity, (6) includes both and shows that the interaction effect of being far from London is attributable to the model-implied elasticity. In (1) and (4)–(6) an intercept is not shown. Standard errors clustered by district in (1)–(3) or robust standard errors in (4)–(6) in parentheses.

Table I.1: Relationship between industrialization and cultural change, moderated by model-implied elasticities

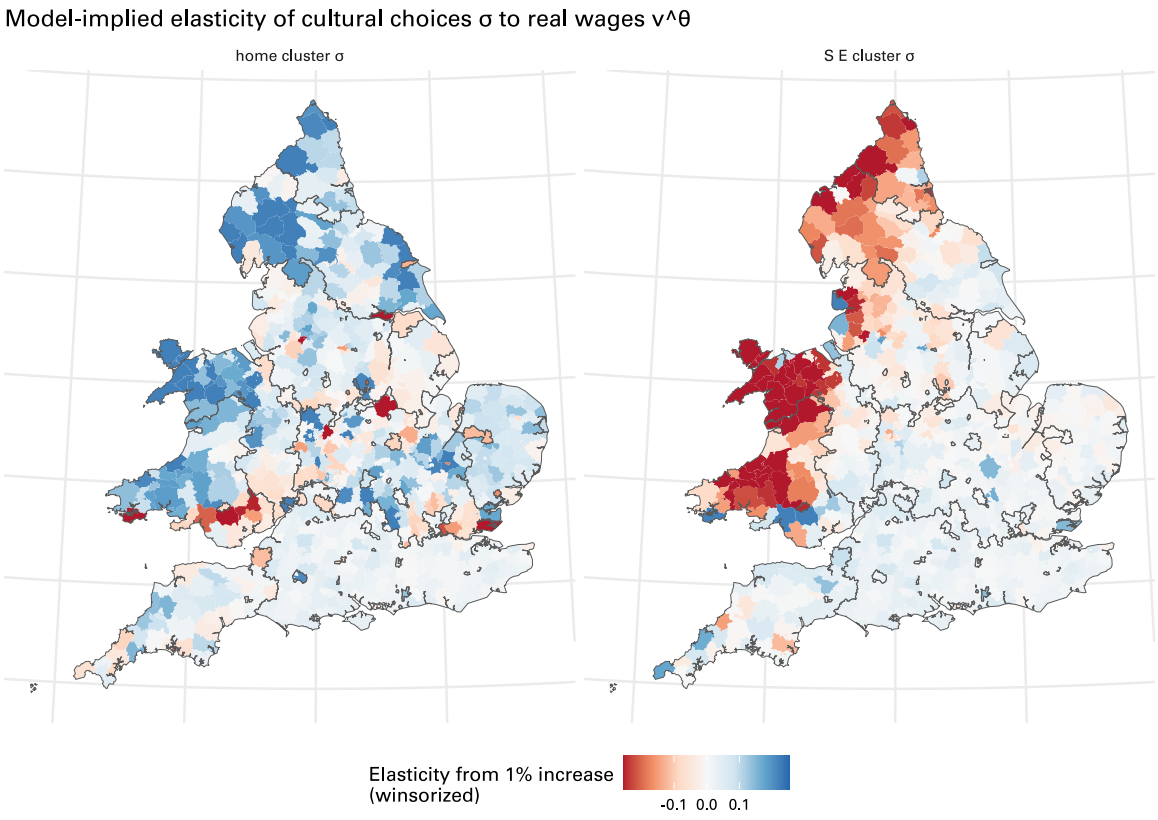


Figure I.1: Elasticities of cultural choices to real wages

We run one counterfactual simulation for each district in which we increase the real wage v_d^θ by 1%, and solve for the new equilibrium cultural choices σ in each location. The map on the left shows the elasticity implied by this 1% shock of σ for each location's home culture; the map on the right shows the elasticity of σ for the Southeast culture.

Note that the term on the right is a summation over all different cultures, and so will feature in elasticities of cultural choice to economic change regardless of the culture in question. Within-location variation in how individuals embrace different cultures in response to economic change will depend only on the left term, the elasticity of Ω_o^k to v_o^θ. The real wage v_o^θ enters Ω_o^k both directly, and by affecting the number of members of culture k in o through migration. Taking the log partial derivative of Ω_o^k with respect to v_o^θ, holding m_o^k constant, gives the following expression:

$$\frac{\partial \ln \Omega_o^k}{\partial \ln v_o^\theta} = \frac{\left(v_o \delta_{oo} \xi_o^k\right)^\theta \left(m_o^k\right)^\alpha}{\Omega_o^k}. \tag{8}$$

This partial equilibrium elasticity is equivalent to the share of people of culture k in location o who choose to remain in location o. We calculate this quantity using the gravity fixed effects (which provide $\left(v_o \xi_o^k\right)^\theta \left(m_o^k\right)^\alpha$), estimated migration costs, and calculated values of Ω_o^k from Section 4. Conditional on district fixed effects, this measure is very strongly correlated with the general equilibrium elasticities (Table I.2), with an R^2 of 0.92.

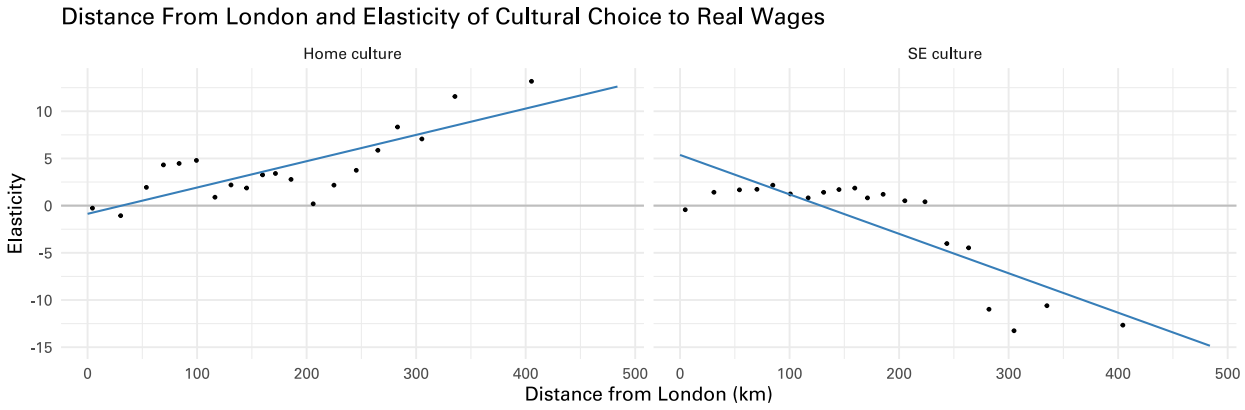


Figure I.2: Elasticity of cultural choices to real wages, relative to distance to London

This figure plots the general equilibrium elasticity of the home culture (left) and Southeast English culture (right) to the district real wage, against distance to London.

Equation (8) helps us say more about which cultures in which locations will grow in response to local economic growth, and why. The culture-specific components of the numerator capture exogenous (ξ_o^k) and endogenous (m_o^k) preferences of people of culture k for that location. The denominator (Ω_o^k) captures these culture-specific preferences for other locations, weighted by real wages and migration costs. A positive shock to the real wage in location o influences people's choices for culture k more than other cultures if there are stronger non-economic benefits to being of culture k if one lives in o than in other locations, relative to that locational tradeoff for other cultures.

The lower rates of migration to and from peripheral regions account for the core-periphery pattern in home culture elasticities. Figure I.3A plots the partial equilibrium elasticity for the home and other cultures, against distance from London. Going further from London, the home culture elasticity increases more, indicating that a positive shock to wages increases the proportion choosing the home culture. Panels B and C decompose the contributions of culture-specific populations (m_d^k) and cultural amenities (ξ_d^k), by calculating the equivalent of the partial equilibrium elasticity excluding cultural amenities (in B) or population (in C). These figures show that culture-specific patterns of migration, not differences in cultural amenities, account for the core-periphery pattern in home culture elasticities. In peripheral regions, distance-related migration costs lead to less out-migration by members of the home culture and less in-migration by other cultures which serves to reinforce the home culture. Examining the partial equilibrium elasticity for the Southeast culture relative to other cultures in Figure I.4 reveals an essentially flipped pattern: locations further from London attract fewer members of the Southeast culture and so have weaker elasticities for that culture relative to others.

	General equilibrium elasticity			
	(1)	(2)	(3)	(4)
Partial equilibrium elasticity	5.650	344.083	5.523	344.234
	(0.763)	(5.424)	(0.755)	(5.307)
District FE		x		x
Cluster FE			x	x
N	7345	7345	7345	7345
R^2	0.011	0.924	0.018	0.925

This table shows OLS estimates at the district-by-cluster level. The dependent variable is the model-implied general equilibrium elasticity of σ_o^k to v_o^θ, the independent variable is the partial equilibrium elasticity, calculated as the gravity-model predicted share of people of culture k in o not migrating from location o. Models (2) and (4) include district fixed effects, (3) and (4) cluster fixed effects. Standard errors clustered by district in parentheses.

Table I.2: Relationship between partial and general equilibrium elasticities

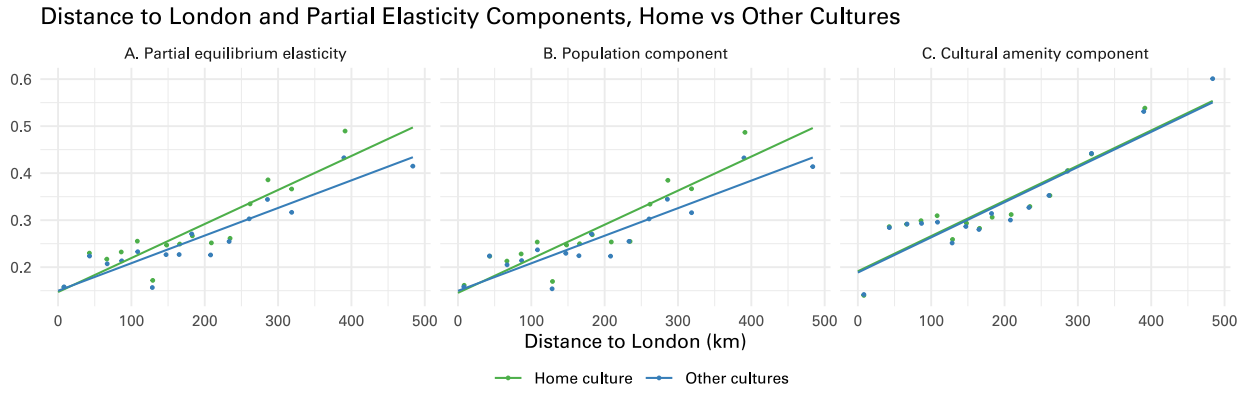


Figure I.3: Decomposing the relationship between distance to London and partial equilibrium elasticities for different cultures

Panel A plots the partial equilibrium elasticity, defined in Equation (8) for the home culture and other cultures, against distance to London. The figure is a binned scatterplot with OLS overlay. The larger gap between the home and other culture values further from London indicates a stronger predicted effect of real wages on choosing the home culture further from London. Panel B plots the *population component*$_o^k := \frac{(v_o\delta_{oo})^\theta \left(m_o^k\right)^\alpha}{\sum_{d=1}^N (v_d\delta_{od})^\theta \left(m_d^k\right)^\alpha}$ against distance to London. This variable corresponds to the partial equilibrium elasticity if cultural amenities were the same everywhere and measures the extent to which the partial equilibrium elasticity is influenced by the culture's population being concentrated in the location in question. Panel C plots the *cultural amenity component*$_o^k := \frac{\left(v_o\delta_{oo}\xi_o^k\right)^\theta}{\sum_{d=1}^N \left(v_d\delta_{od}\xi_d^k\right)^\theta}$ against distance to London. This variable is the partial elasticity if destination populations did not influence migration and cultural choices and measures the extent to which the partial equilibrium elasticity is influenced by amenities for the culture being concentrated in the location in question.

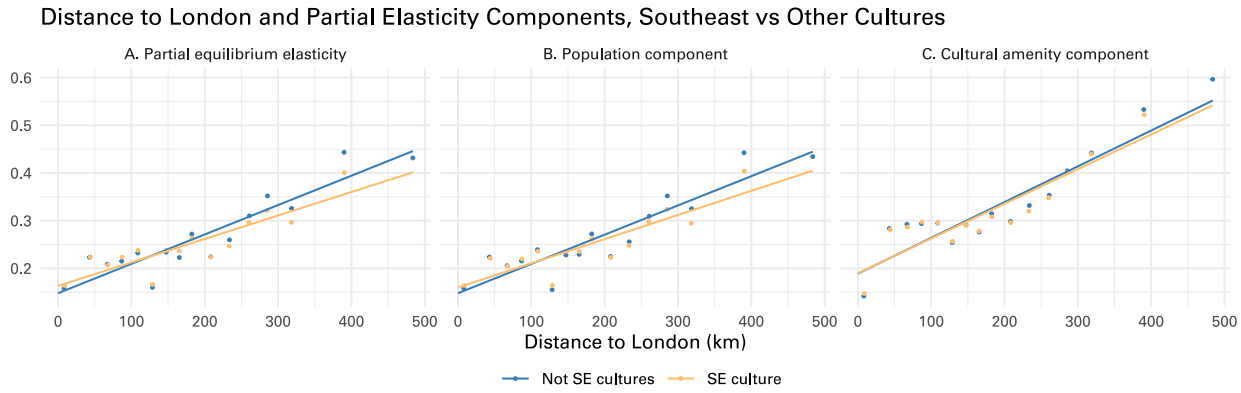


Figure I.4: Decomposing the relationship between distance to London and partial equilibrium elasticities for the Southeast and other cultures

Panel A plots the partial equilibrium elasticity, defined in Equation (8) for the Southeast English culture and other cultures, against distance to London. The figure is a binned scatterplot with OLS overlay. Panel B plots the contribution of population to this pattern, as defined in Figure I.3, C plots the contribution of cultural amenities, also as defined in Figure I.3.

Table J.1: Counterfactual estimates from homogenizing culture

Counterfactual	%Δ Ave. v_d^θ	%Δ share migrant
Removing non-S E ψ_o^k	0.757	0.702
Removing non-S E ψ_o^k and fixing $\xi_d^k = 1$	-0.152	0.249
Doubling S E ψ_o^k	0.755	0.696
Halving S E ψ_o^k	-0.317	-0.015
Fixing ψ_o^k to have same average for all cultures	-0.116	0.144

This table shows percentage changes in the average v_d^θ and share migrating under different counterfactual scenarios. The first replaces all cultural choice preferences ψ_o^k with zero for cultures other than the Southeast. The second adds to this specification by also fixing ξ_d^k to one for all destinations. The third replaces ψ_o^k with double its value for the Southeast culture, the fourth with half its value. The fifth rescales ψ_o^k so that the average value for all cultures is equal.

J ADDITIONAL ANALYSES

J.1 *Effects of Cultural Homogeneity on Migration and Economic Outcomes*

This section examines whether rising cultural homogeneity affected economic outcomes. Because culture-specific preferences and homophily influenced where people migrated, changing the popularity of different cultures should alter the spatial distribution of population, and, if doing so directs more people towards locations with higher real wages, increase average income (v_d). Table J.1 shows the results of different counterfactual simulations, primarily changing the ψ_o^k matrix of place-specific preferences for choosing particular cultures. Total assimilation into the Southeast culture, achieved by setting all non-Southeast ψ_o^k values to zero, increases the average v_d^θ by 0.76%.

If we think of v_d as representing the real wage in location d, under constant returns to scale, then given an estimate of θ, the elasticity of migration to wages, we can infer the effect on average wages. Tombe and Zhu (2019), using data from China, estimate elasticities of migration to wages between 1.2 and 1.6, and use 1.5 in their analysis, which would suggest that the estimate in the first row of Table J.1 corresponds to around a 0.5% increase in average wages. Morten and Oliveira (2023) estimate a migration elasticity of 4.5 in Brazil, corresponding to a 0.2% increase in average wages in response to full homogenization into the Southeast culture. Caliendo et al. (2021) estimate a much lower elasticity of 0.5—corresponding to a 1.5% increase in average wages—albeit in a dynamic model of migration in Europe. The positive effect of homogenization into the Southeast culture on economic output is due to the correlation between cultural amenities (ξ_d^k) specific to the Southeast culture and real wages: increasing the prevalence of the Southeast culture directs more people to high real wage locations. If we set all culture-specific destination migration preferences to 1, we find that cultural homogeneity in fact decreased average v_d^θ. Figure J.1 shows that both counterfactuals redistribute population across regions, away from rural parts of the North and Wales. More moderate changes to the popularity of the Southeast cluster have more moderate effects, while rescaling ψ_o^k so that the average value for all cultures is equal decreases average v_d^θ.

These small effects of homogenization on economic output make sense given the paper's other findings. Distance-related migration costs that help preserve peripheral cultures by diminishing out-migration also keep workers in low-productivity peripheral regions, even in

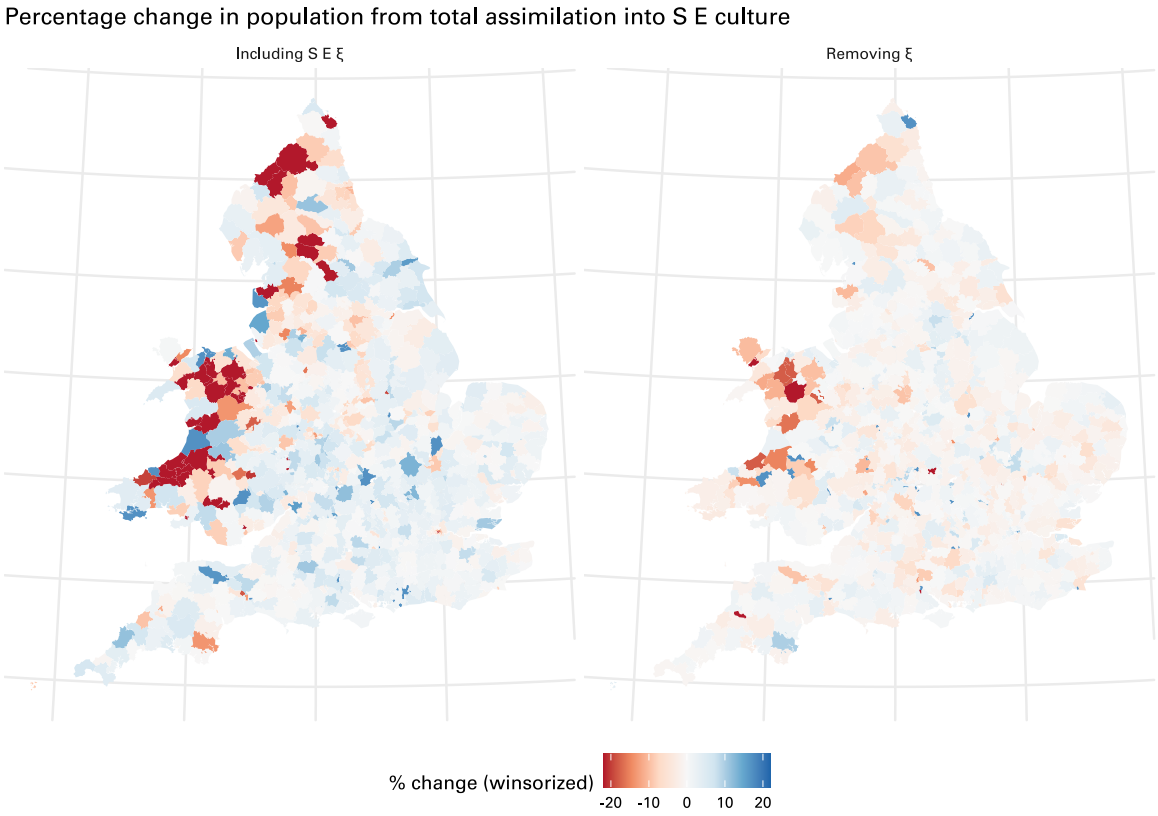


Figure J.1: Changes in population distribution from homogenizing culture

These maps show the percentage change in destination populations under counterfactual simulations that replace cultural preference terms ψ_o^k for all cultures except the Southeast culture with zero. The figure on the right also replaces all culture-specific migration preferences $\left(\xi_d^k\right)^\theta$ with 1.

the absence of cultural sorting and preferences. The model already assumes that individuals migrate towards higher-wage locations, which limits the extent to which eliminating other factors can shift migration towards such locations.

J.2 *Measuring Transportation Costs*

In the main version of the model, transportation costs are simply a function of geographic distance and an indicator that the origin is the same as the destination. In this section we explore directly calculating transportation costs using shapefiles for railways, stations, shipping routes, and ports.

We use data on railways in 1851 and 1881 from Satchell et al. (2023) and Marti-Henneberg et al. (2023), stations in 1851 and 1881 from Marti-Henneberg et al. (2018) and Henneberg et al. (2018), and ports and shipping routes from Alvarez-Palau et al. (2019).

Our method closely follows Donaldson and Hornbeck (2016) in constructing a matrix of transportation costs between Registration District Centroids. With this data for 1851 and 1881, we create a network database in which nodes are locations—district centroids, stations, and ports—and edges are railways, shipping routes, and land routes connecting them. In

addition to railways and shipping we allow any two nodes within 100km of each other to be connected by land. We weight these edges using the transportation cost parameters from Donaldson and Hornbeck (2016) and then use Dijkstra's algorithm to calculate the least cost path between any two district centroids. One should note that Donaldson and Hornbeck are studying trade costs in the US during the same period, and so it is possible that the relative cost of different modes of transport for shipping goods differs from the relative cost for transporting people.

Figure J.2 plots these calculated transport costs against distance in 1851 and 1881. As expected, transportation costs strongly correlate with geographic distance. The expansion of the railway network served to reduce transportation costs at intermediate-level distances.

Table J.2 shows the relationship between these transportation costs and migration flows. These models regress log migration flows against log transportation costs, with origin-year and destination-year fixed effects. In the context of the gravity model of migration, one can think of the resulting estimates as indicating the relationship between transportation costs and the migration costs implied by migration flows. In model (1) we find a strong negative relationship between transportation costs and migration flows, consistent with the idea that transportation costs influenced migration costs. However, in (2) we find that controlling for geographic distance shrink the coefficient on transportation costs to zero. The implication is that (1) picks up the effects of distance and not a distinct effect of infrastructure. In (3) we add origin-destination fixed effects and similarly find a null effect of transportation costs on migration flows. Note that in all these specifications, we would expect the coefficient on transportation costs to be biased away from zero in the negative direction due to the endogeneity of infrastructure construction: demand for migration from an origin to a destination could drive investment in connections between the two locations.

Our conclusion from this exercise is that geographic distance provides a better measure of migration costs than measures of transportation costs. In addition to possible error in the measurement of transportation costs, it is plausible that distance captures something related to ease of access or familiarity that is distinct from the physical cost of transportation.

J.3 *Relationship Between Amenities and Wages*

Do cultural amenities capture just taste, or also local wage differentials? If members of a given culture are paid more than members of other cultures in a given location, that would induce members of that culture in particular to migrate there, and would be picked up by the cultural amenity parameter in our model. Such a wage difference could emerge if there is demand for specific skills in a location, which members of some cultures but not others tend to have, or if there is taste-based discrimination against members of certain cultures in the labor market.

Table J.3 examines this claim, using HISCAM scores (Lambert et al., 2013). These provide a measure of the socioeconomic status of each occupation. Our dependent variable is the average HISCAM score, at the culture-district level, for men born 1861–1895. A higher value implies that men from that culture living in that location tended to have higher status occupation. Our data does not allow us to examine whether men were paid more to do the same job. We regress average HISCAM scores against real wages (v_d^θ), cultural amenities (ξ_d^k), and transmission tastes (ψ_o^k), weighting by the number of individuals in each cell.

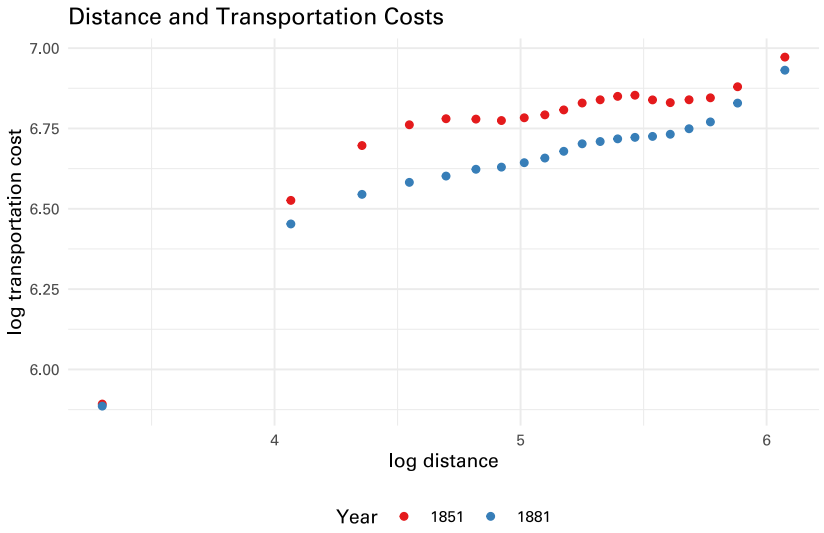


Figure J.2: Distance and Transport Costs, 1851 and 1881

Binned scatterplot of log transportation costs (y axis) in 1851 and 1881 against log distance.

This specification gives equivalent coefficients to an individual-level regression. We find an ambiguous relationship between cultural amenities and transmission tastes, and HISCAM scores: in two specifications the taste parameter is positively correlated with status, in two they are negatively correlated, though only the negative coefficients are statistically significant. These results provides evidence against the idea that cultural amenities capture positive local wage differentials.

A long tradition in urban economics dating back to Roback (1982) argues that amenities that affect workers' preferences but not firm productivity should lead to lower wages. The logic is that there is a compensating differential: if in equilibrium the marginal worker is indifferent across locations, a factor that makes a location more desirable will attract more workers and push down wages. While we do not explicitly model a culture-by-location process that would generate such a compensating differential, Bryan and Morten (2019) suggest one way to reconcile a negative relationship between cultural amenities and wages with our theoretical and empirical framework. In their model of migration, the individual-by-destination Frèchet shock (the equivalent of ε_{id} in our model) is interpreted as a location-specific skill. A higher amenity value for members of a given culture in a given location serves to attract members of that culture with lower Frèchet shocks for that location, who have less-valuable location-specific skills and earn less.

J.4 *Accounting for Emigration*

During this period, vast numbers of English and Welsh people emigrated, especially to the US, Canada, Australia, and New Zealand. These emigrants are not in our dataset, and to the best of our knowledge, there is no comprehensive data on emigrants. The question is how emigration affects our results. Here we show that our theory and evidence related to internal migration is unaffected by the existence of emigration. The potential for emigration

	log migrants		
	(1)	(2)	(3)
log transportation cost	−2.422	0.105	−0.089
	(0.098)	(0.081)	(0.060)
log distance		−1.743	
		(0.071)	
Origin-year FE	x	x	x
Destination-year FE	x	x	x
Origin-destination FE			x
N	1382742	1382742	795729

This table presents evidence of the relationship between transportation cost, distance, and migration flows in the 1851 and 1881 censuses. Models are estimated by Poisson Pseudo-Maximum Likelihood. Data is at the origin-destination-year level. The dependent variable is the log number migrating from the origin to the destination in that year's census. The independent variable of interest is the log transportation cost. All models include origin-year and destination-year fixed effects. Model (2) controls for the log distance between the origin and destination. (3) adds origin-destination fixed effects, which are collinear with distance. Standard errors clustered by origin and destination in parentheses.

Table J.2: Null relationship between transport costs and migration

	Average economic status			
	(1)	(2)	(3)	(4)
log real wage (v_d^θ)	0.312		0.283	
	(0.364)		(0.368)	
log cultural amenity (ξ)	−5.109	0.332		
	(1.285)	(0.303)		
log cultural transmission taste (ψ)			0.733	−0.809
			(0.439)	(0.122)
Cluster FE	x	x	x	x
District FE		x		x
N	7385	7385	7375	7375
R^2	0.080	0.979	0.071	0.981

This table shows OLS estimates at the district-cluster level. The dependent variable is the average economic status inferred using the HISCAM score corresponding to each occupation, for men born 1861–1895 resident in the district, allocated to clusters by their first names. In models (1) and (2) the independent variable is the destination (log) cultural amenity (ξ), in (3) and (4) the (log) cultural transmission taste (ψ). All models include cluster fixed effects. Models (1) and (3) also include the destination (log) real wage (v_d^θ), (2) and (4) add district fixed effects which are collinear with real wages. Observations are weighted by the number of individuals for which occupational status data is available Standard errors clustered by district in parentheses.

Table J.3: Relationship between cultural amenities, real wages, transmission tastes, and economic status

does affect cultural choices, but in our framework it is captured by the location-by-culture cultural transmission tastes. Our counterfactual exercises must be interpreted as holding the ratio of domestic to international migration opportunities fixed. This assumption is appropriate given we are interested in examining how changing patterns of industrialization and internal migration affect identity choice, not in how changing international opportunities affect cultural choices.

Suppose that in addition to the locations we observe, there are a number of unobserved locations, to which people can also migrate. In the main body of the paper, locations 1 through N are domestic. Let us add $F - N$ foreign locations, so that there are F total locations, where 1 through N are domestic and $N + 1$ through F are foreign.

Denote the number of people born in o assigned to culture k, including those who emigrate, with $n_o^k + e_o^k$, where e_o^k refers to the number of emigrants from o of culture k, and n_o^k, as in our analysis, is the number originating in o of culture k who remain in England and Wales. Then the number migrating to each location d from domestic locations (Equation (1)) becomes

$$m_{od}^k = \frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\sum_{f=1}^F \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha} \left(n_o^k + e_o^k\right). \tag{9}$$

Note that the number from o of culture k who remain in England and Wales, n_o^k, is the sum over domestic locations:

$$n_o^k = \sum_{j=1}^N m_{oj}^k = \frac{\sum_{j=1}^N \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha}{\sum_{f=1}^F \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha} \left(n_o^k + e_o^k\right),$$

which we rearrange as follows

$$\frac{n_o^k}{n_o^k + e_o^k} \frac{\sum_{f=1}^F \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha}{\sum_{j=1}^N \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha} = 1.$$

Multiplying (9) by the left-hand side of this equation gives

$$m_{od}^k = \underbrace{\frac{\left(v_d \delta_{od} \xi_d^k\right)^\theta \left(m_d^k\right)^\alpha}{\sum_{j=1}^N \left(v_j \delta_{oj} \xi_j^k\right)^\theta \left(m_j^k\right)^\alpha} n_o^k}_{m_{od}^k \text{ in (1)}} \underbrace{\frac{\sum_{f=1}^F \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha}{\sum_{f=1}^F \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha} \frac{n_o^k + e_o^k}{n_o^k + e_o^k}}_{=1}.$$

This exercise shows that allowing for emigration and unobserved destination locations does not affect the part of our model that analyzes migration choices and so does not affect our calculations of the relevant v_d, δ_{od} and ξ_d^k components. The logic is that we are already implicitly conditioning on people remaining in England and Wales.

Counterfactuals that set wages back to 1850s levels should affect emigration, for these analyses one must in effect assume that we are holding the ratio of domestic to foreign migration opportunities fixed.

Including foreign locations does alter the expressions for cultural choices. If foreign

migration opportunities also matter for cultural choices, then in addition to the domestic Ω_o^k, there is an additional foreign term:

$$\Omega_o^{k,t} := \sum_{f=1}^{F} \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha = \Omega_o^k + \Omega_o^{k,f}.$$

Here $\Omega_o^{k,t}$ is the culture-specific expected utility inclusive of emigration, and

$$\Omega_o^{k,f} := \sum_{f=N+1}^{F} \left(v_f \delta_{of} \xi_f^k\right)^\theta \left(m_f^k\right)^\alpha$$

is the expected utility just from foreign locations. Modifying the expression for cultural choices in (3) to include foreign migration opportunities gives

$$\sigma_o^k = \frac{\left(\Omega_o^{k,t}\right)^\varphi \psi_o^k}{\sum_{l=1}^{K} \left(\Omega_o^{l,t}\right)^\varphi \psi_l^k}.$$

Defining the ratio of total to domestic culture-specific expected utility as $\omega_o^k := \frac{\Omega_o^{k,t}}{\Omega_o^k}$, we have

$$\sigma_o^k = \frac{\left(\Omega_o^k\right)^\varphi \left(\omega_o^k\right)^\varphi \psi_o^k}{\sum_{l=1}^{K} \left(\Omega_o^l\right)^\varphi \left(\omega_o^l\right)^\varphi \psi_l^k}.$$

Taking logarithms as in (5) to create an estimating equation, we see that the error term in this regression now includes the ω_o^k term:

$$\ln \sigma_o^k = \varphi \ln \Omega_o^k - \underbrace{\ln \left(\sum_{l=1}^{K} \left(\Omega_o^l\right)^\varphi \left(\omega_o^l\right)^\varphi \psi_l^k \right)}_{\text{place FE}} + \underbrace{\varphi \ln \omega_o^k + \ln \psi_o^k}_{\text{Error}}.$$

The implication is that our estimates of the cultural transmission taste ψ_o^k also include the ratio of total to domestic migration opportunities, ω_o^k. If it were easier to practice one culture in the United States than others, then locations with greater ease of migration to the United States would appear to have a stronger cultural transmission taste for that culture. Our counterfactuals should then be interpreted as providing the estimated effects of changes in the domestic economy, holding fixed the ratio of domestic to foreign opportunities.

J.5 *Alternative Elasticities*

A key result in this paper is that the change in cultural choices going from a counterfactual estimated using 1851 economic fundamentals to the observed 1911 data matches specific observed cultural changes. We show that the culture of the Southeast of England expands, especially in places further from London, while peripheral home cultures decline, an effect moderated by the presence of coal. Here we investigate whether the ability of our model

to reproduce these patterns is sensitive to the estimated homophily (α) and culture (φ) elasticities.

For an arbitrary pair of elasticities, we back out the other parameters of the model, and solve for counterfactual cultural choices under 1851 real wages, populations, and migration costs. Figure J.3 plots the equivalent of Figure 8A at each of these pairs of elasticities, Figure J.4 plots the equivalent of Figure 9A. Note that in both figures the scale of the y axis varies across rows. We only include pairs of elasticities such that $\alpha\varphi < 1$, because for values above that bound the equilibrium is not unique and so simulation results are hard to interpret. The broad takeaway is that the qualitative predictions of an increase in the Southeast culture further from London, and a decline in the home culture further from London moderated by the presence of coal hold across non-zero values of these elasticities. However, the magnitude of the predicted changes, and the steepness of the slope linking distance to London and predicted changes, increases considerably as we increase the magnitude of the elasticities.

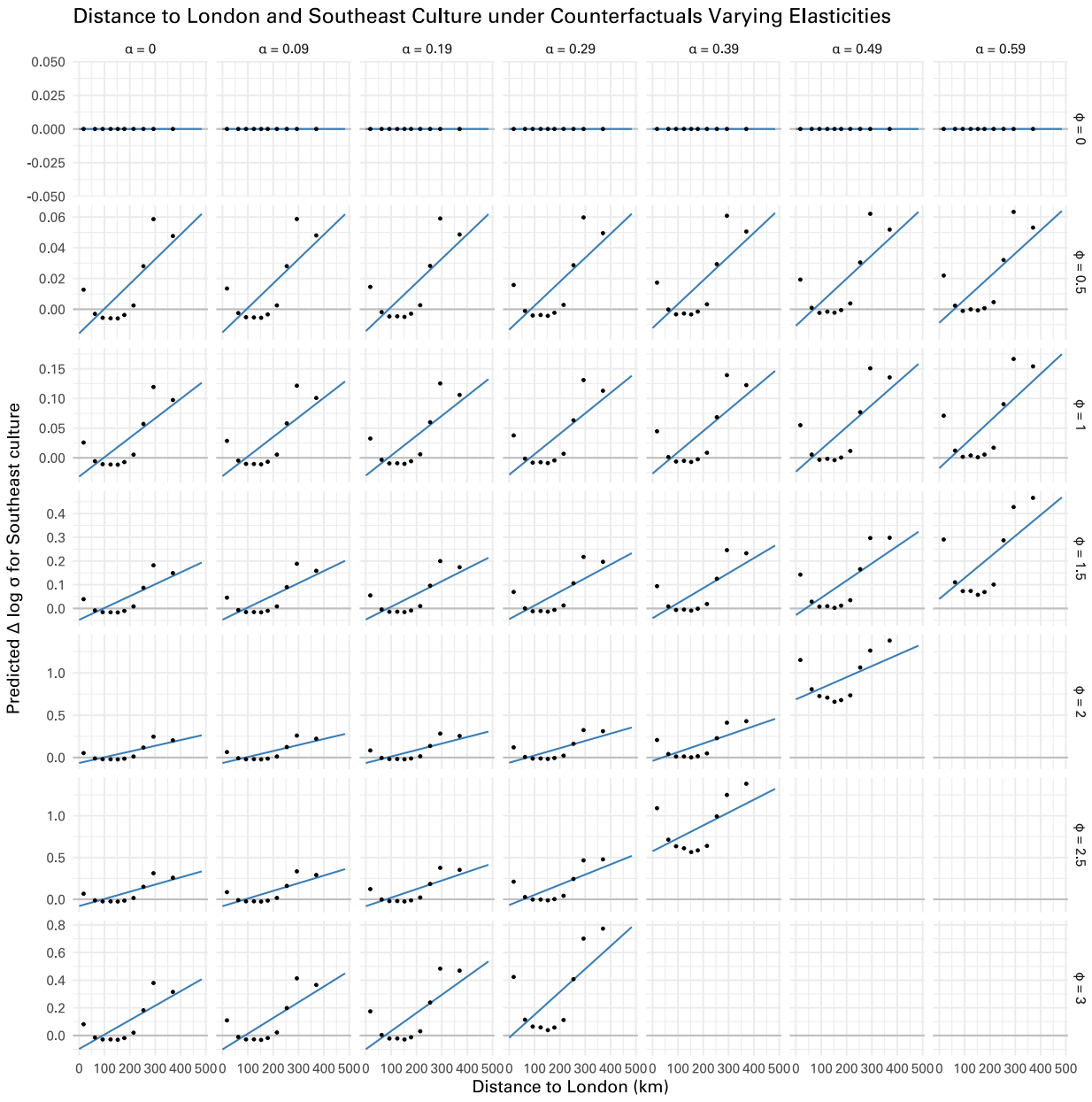


Figure J.3: Relationship between distance to London and the rise of the Southeast culture under different elasticities

Each grid cell shows the binned scatter relationship between distance to London and the log share choosing the Southeast culture in the observed data minus that in a counterfactual estimated using 1851 real wages, populations and migration costs. Each cell varies the homophily (α) and culture elasticities (φ). Instead of estimating these elasticities using instrumental variables, we fix them at a given value and proceed to back out the other parameters of the data, then solve for counterfactuals using these assumed elasticities and backed out parameters. Columns vary α, rows φ. Note that the scale of the y axis varies across rows. We do not calculate counterfactuals for values of $\alpha\varphi \geq 1$ as the equilibrium is non-unique and unstable.

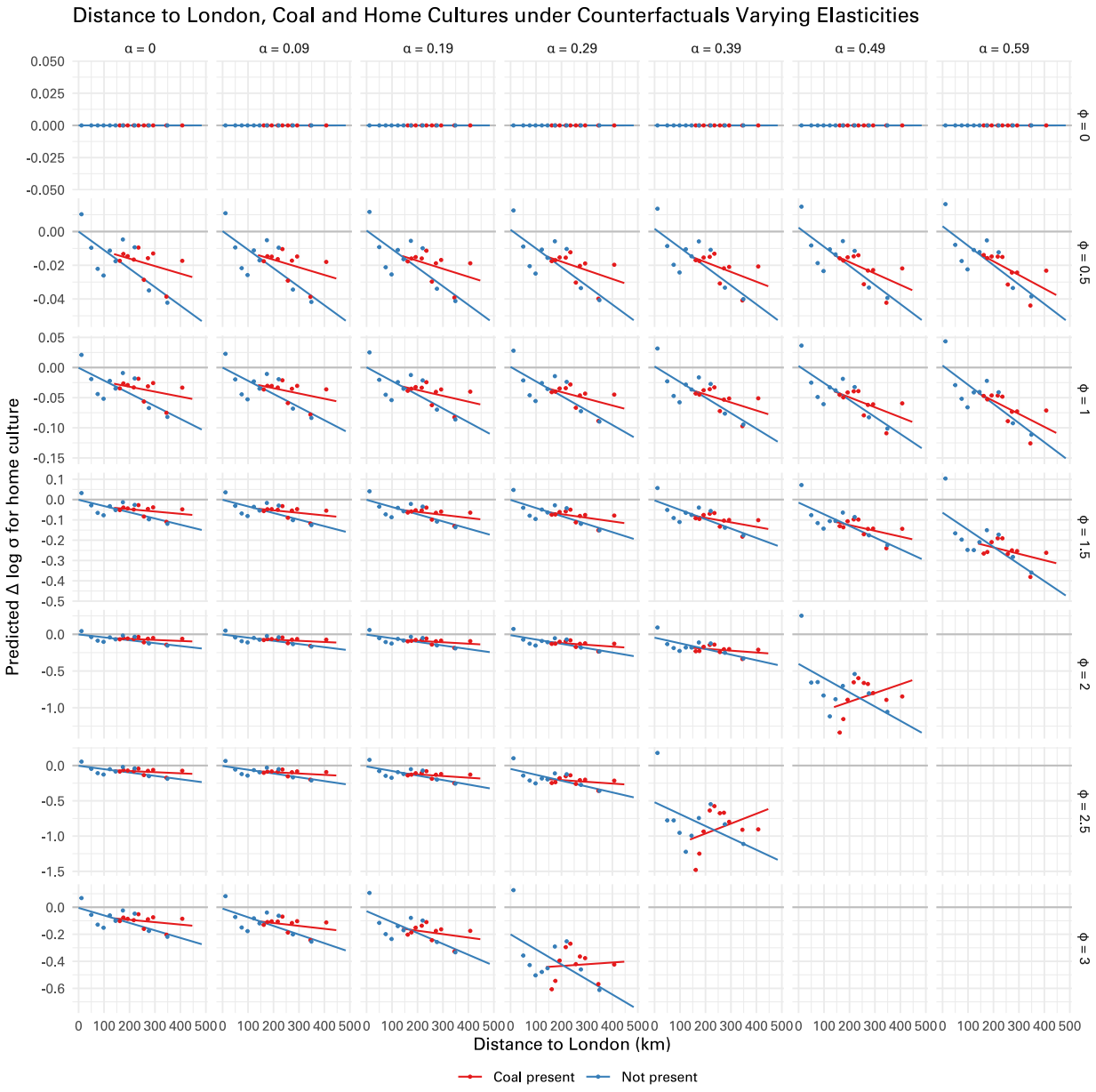


Figure J.4: Relationship between distance to London, coal, and the decline of the home culture under different elasticities

Each grid cell shows the binned scatter relationship between distance to London and the log share choosing the home culture in the observed data minus that in a counterfactual estimated using 1851 real wages, populations and migration costs, subset by whether the district contains coal. Each cell varies the homophily (α) and culture elasticities (φ). Instead of estimating these elasticities using instrumental variables, we fix them at a given value and proceed to back out the other parameters of the data, then solve for counterfactuals using these assumed elasticities and backed out parameters. Columns vary α, rows φ. Note that the scale of the y axis varies across rows. We do not calculate counterfactuals for values of $\alpha\varphi \geq 1$ as the equilibrium is non-unique and unstable.

J.6 *Alternative Cultural Clusters*

In this section we examine how results from the paper change using alternative cultural clusters. We focus on two sets of clusters. First, we group counties into the regions used to coordinate civil defence during the Second World War. This grouping was made after the changes we study, but it was one of the first administrative geographies above the county level, and gives a similar number of regions to the cultural clusters we estimate. The point is not that this grouping combines culturally-similar regions, but rather provides a way of combining adjacent counties. The left panel of Figure J.5 shows these regions.

Second, we estimate clusters using data on medieval English from *A Linguistic Atlas of Late Medieval English* (Benskin et al., 2013). This source codes the appearance of linguistic features in geocoded medieval texts. We create a measure of the similarity between these sources based on the share of common linguistic features, and then run the Louvain graphical clustering algorithm on the similarity matrix. Relative to the spectral algorithm we use with surname data, the Louvain algorithm is well-suited to data that already encodes the distance between entries. It also automatically estimates the optimal number of clusters. We allocate districts to clusters based on the most common cluster among the 7 sources closest to the district. Because the data only applies to England, we allocate Wales to its own cluster. The right panel of Figure J.5 shows these clusters.

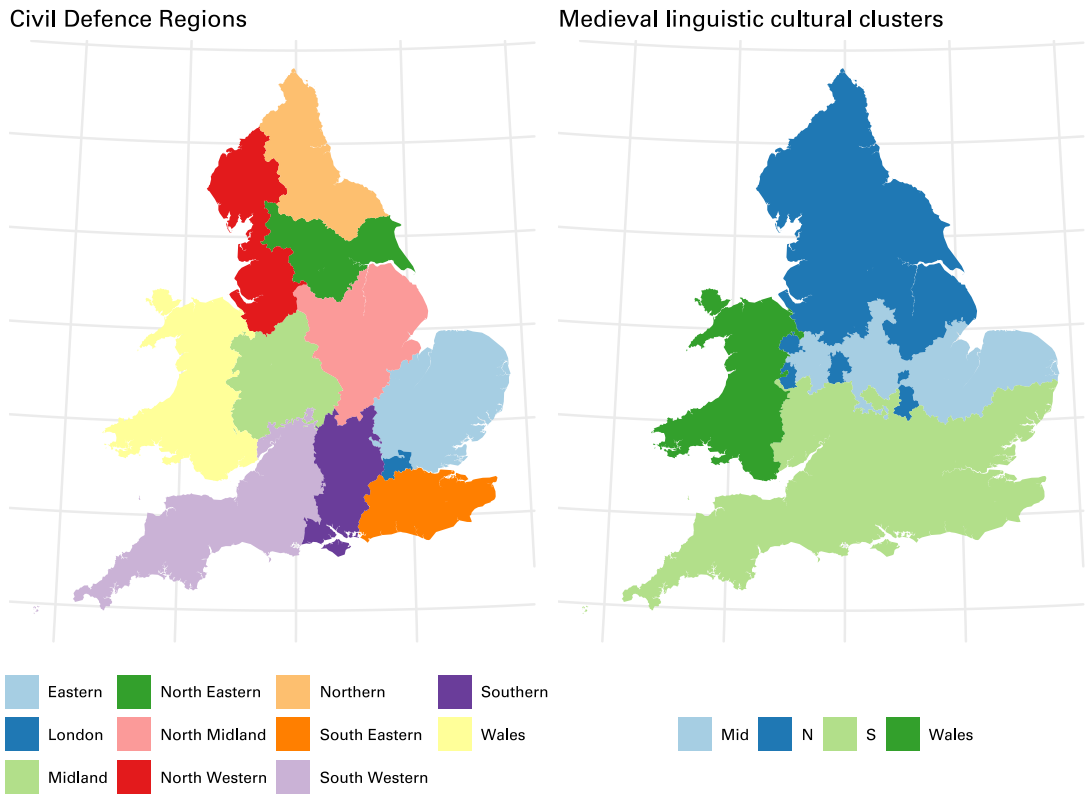


Figure J.5: Alternative cultural clusters

The pattern we observe in Figure 4, whereby by the start of the twentieth century, most of England was sucked into the Southeast-English culture, with exceptions in Wales and the Northeast, holds under these different clusters. Figures J.6 and J.7 both show the cluster with the highest name score among those born in each district 1851–1860 and 1901–1910, calculated using the two alternative clusters. Using civil defence regions, the hegemonic culture by 1910 is London, with linguistic cultures, the large Southern English cluster. These patterns suggest that the rise of the Southeast that we observe in the main text is primarily the rise of London.

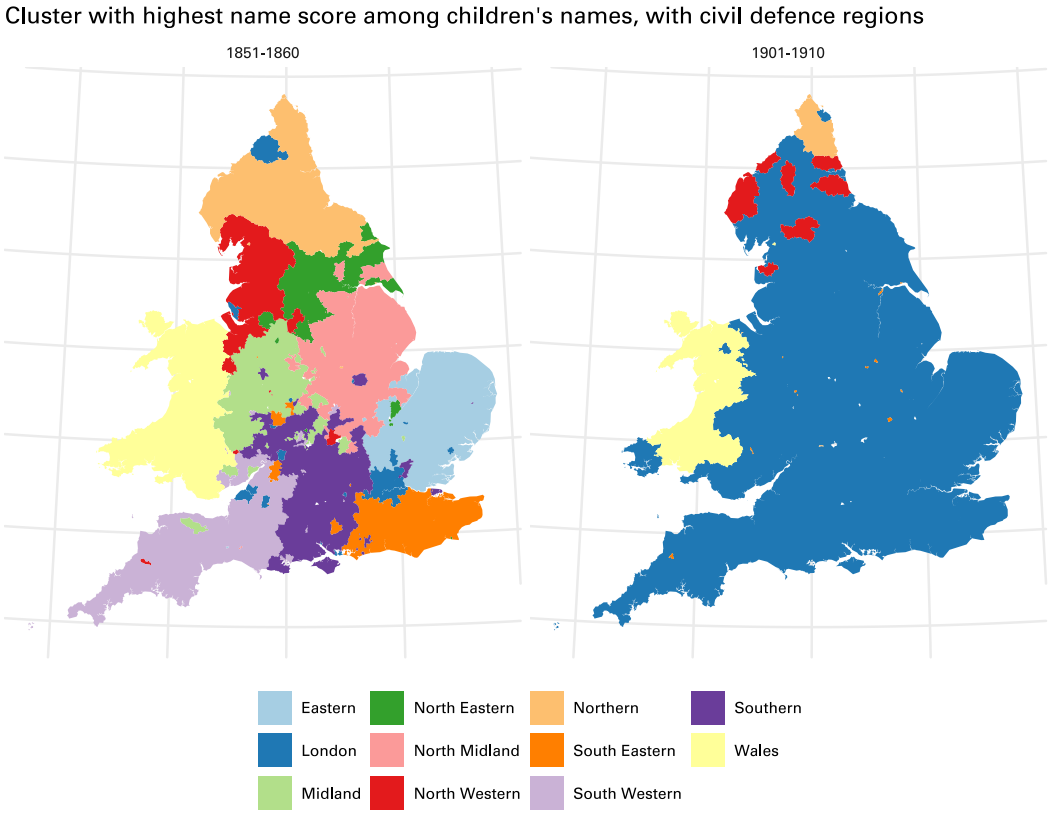


Figure J.6: Changing popularity of cultures, using civil defence regions as clusters

We examine how both our findings change as we change the cultural clusters. We calculate name scores and cluster-by-migration flows using the new clusters, and calibrate the model using the elasticities we estimate in the main body of the paper. We then examine whether the cultural changes predicted by the model conform to actual cultural changes, as in Table 4. We re-run the main counterfactual simulations and examine the new predictions for the aggregate changes of the hegemonic culture (London in the case of the civil defence regions, and the large Southern cluster in the case of the linguistic clusters).

Broadly speaking, our results with these alternative clusters are similar to those using clusters based on historical surnames. Tables J.4 and J.5 show the relationship between changes in predicted and actual cultural choices. The results for the civil defence regions are very close to our baseline estimates. The relationship between the model's predicted cultural

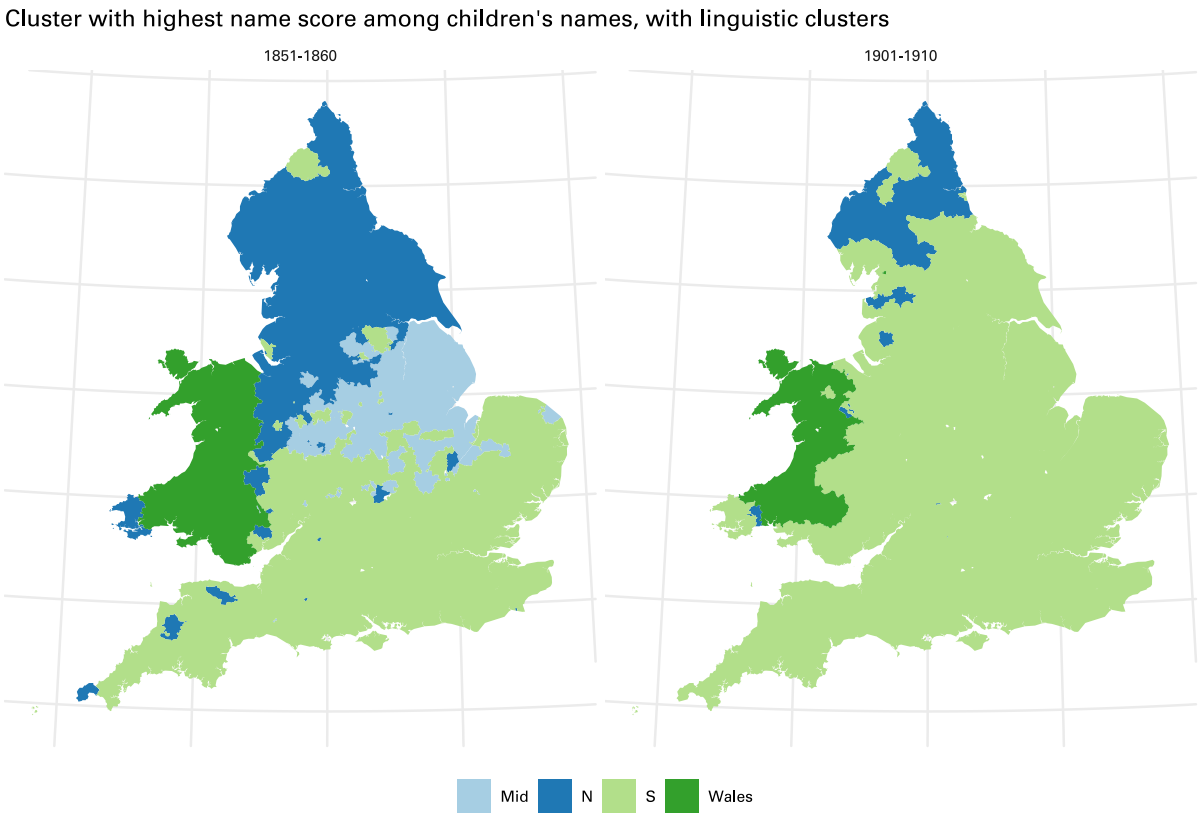


Figure J.7: Changing popularity of cultures, using linguistic clusters

changes and observed cultural changes, with linguistic clusters and both district and cluster fixed effects in Table J.5, is not robust (though the other coefficients in that table are positive and precisely-estimated). This is likely attributable to there being fewer linguistic clusters, which means there is less within-district variation.

In the structural analyses, the different-sized clusters lead to differences in quantitative magnitudes, and to the effects on the home culture, but other effects are qualitatively similar. Tables J.6 and J.7 show the same counterfactual outputs as Table 5, using the alternative clusters. The effects on the popularity of the Southeastern, Southern, and London cultures, are very similar across these specifications, which increases our confidence in our conclusion that economic changes during the 1851–1911 period bolstered the hegemonic culture. The spatial patterns of counterfactual changes in adoption of these cultures are also similar (Figures J.8 and J.9). Our estimates for the effects on the popularity of home cultures differ in their aggregate magnitudes but follow similar spatial patterns (Figures J.10 and Figures J.11).

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.290	0.278	0.166	0.065
	(0.011)	(0.012)	(0.014)	(0.013)
District FE		x		x
Cluster FE			x	x
N	8999	8999	8999	8999
R^2	0.077	0.111	0.645	0.686

This table shows OLS estimates at the district-by-cluster level, replicating Table 4 using civil defence regions. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table J.4: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort, with civil defence regions

	Observed cultural change ($\Delta \ln \sigma$)			
	(1)	(2)	(3)	(4)
Predicted cultural change ($\Delta \ln \sigma$)	0.549	0.609	0.085	−0.070
	(0.025)	(0.027)	(0.013)	(0.012)
District FE		x		x
Cluster FE			x	x
N	3306	3306	3306	3306
R^2	0.199	0.247	0.802	0.867

This table shows OLS estimates at the district-by-cluster level, replicating Table 4 using linguistic clusters. The independent variable is the change in the log share choosing each culture between the counterfactual estimated using 1851 destination real wages v_d^θ, starting populations, and migration costs and the observed value for those born 1861–1895. The dependent variable is the change between the observed value for those born 1841–1860 and those born 1861–1895. Model (2) adds district fixed effects, (3) adds cluster fixed effects, (4) adds both. Standard errors clustered by district in parentheses.

Table J.5: Relationship between the change in log cultural choice shares, σ, relative to the 1851 counterfactual and relative to the 1841–1860 cohort, with linguistic clusters

Table J.6: Counterfactual estimates with civil defence regions

	Home culture pop.		London culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	5.926	7.309	-43.122	-40.602	-35.698	-31.062
1851 real wages, v_d^θ	-5.401	-5.401	-42.189	-42.189	-1.561	-1.561
1851 starting populations	-13.544	-5.174	-29.714	-22.306	-0.727	0.580
1851 migration costs	22.241	22.241	-3.300	-3.300	-31.383	-31.383

This table shows the same counterfactual outputs as Table 5, using civil defence regions in place of clusters based on surnames before 1800.

Table J.7: Counterfactual estimates with linguistic clusters

	Home culture pop.		S culture pop.		Share migrant	
Counterfactual	%Δ	fixing pop.	%Δ	fixing pop.	%Δ	fixing pop.
All 1851 parameters	3.668	-2.359	-20.859	-18.306	-35.218	-30.776
1851 real wages, v_d^θ	-9.744	-9.744	-26.941	-26.941	-1.597	-1.597
1851 starting populations	-1.544	-4.509	-14.534	-7.274	-0.690	0.605
1851 migration costs	10.277	10.277	-0.328	-0.328	-31.099	-31.099

This table shows the same counterfactual outputs as Table 5, using linguistic clusters in place of clusters based on surnames before 1800.

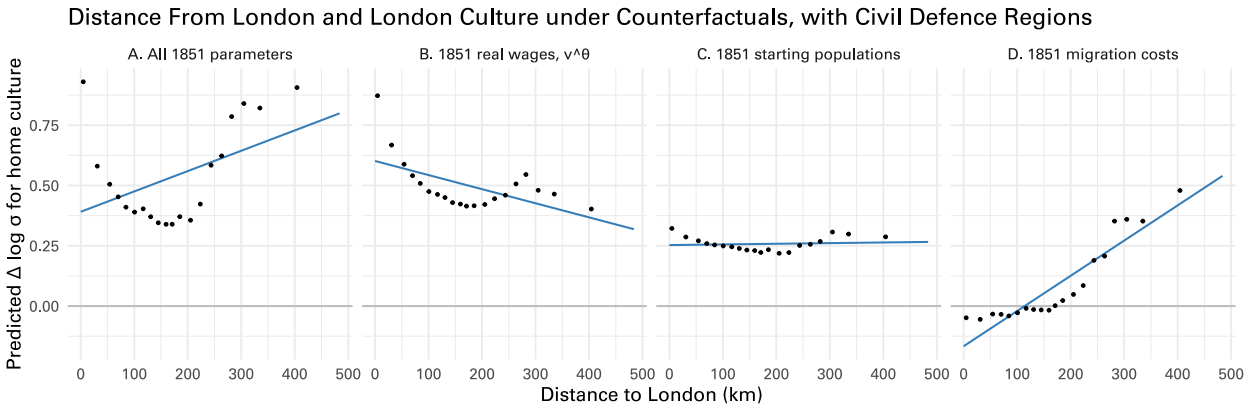


Figure J.8: Distance to London, coal, and the changing popularity of the London culture over the late 19th century, with Civil Defence Regions

This figure replicates Figure 8 using Civil Defence Regions as cultural clusters, and the London region in place of the Southeast region. The figure shows the predicted change in the log share allocated names most associated with the London culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance from the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

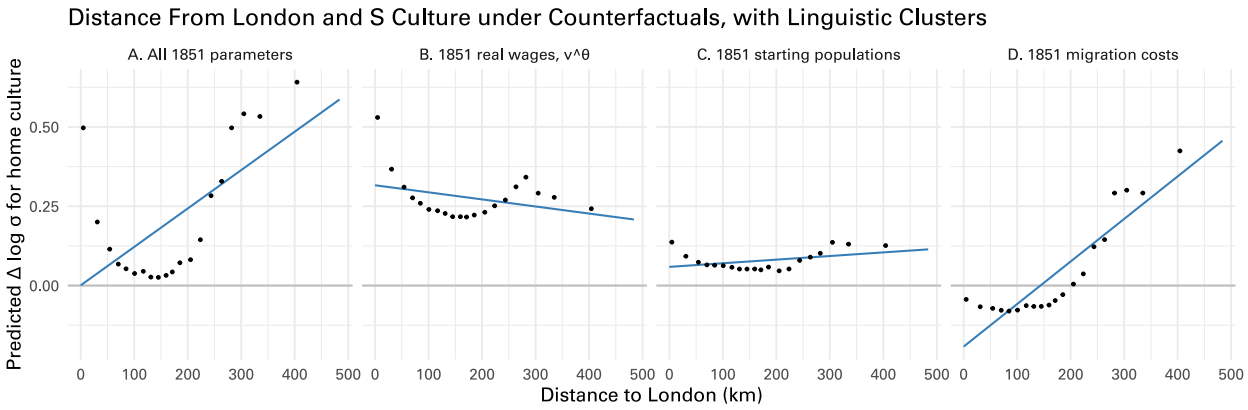


Figure J.9: Distance to London, coal, and the changing popularity of the South English culture over the late 19th century, with linguistic clusters

This figure replicates Figure 8 using linguistic clusters, and the South English cluster, containing London, in place of the Southeast English cluster. The figure shows the predicted change in the log share allocated names most associated with the Southern culture, across different counterfactual scenarios. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

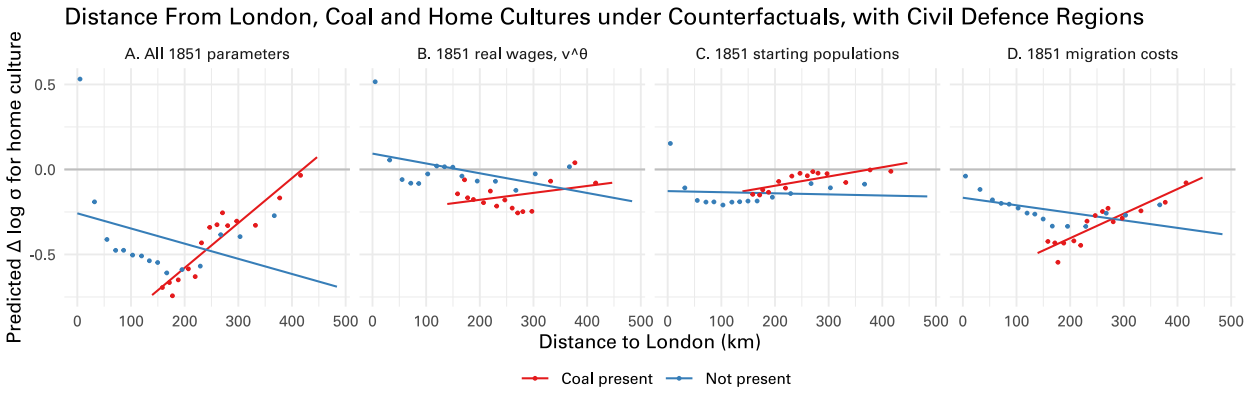


Figure J.10: Average distance, coal, and the changing popularity of home cultures over the late 19th century, with Civil Defence Regions

This figure replicates Figure 9 using Civil Defence Regions as cultural clusters. The figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.

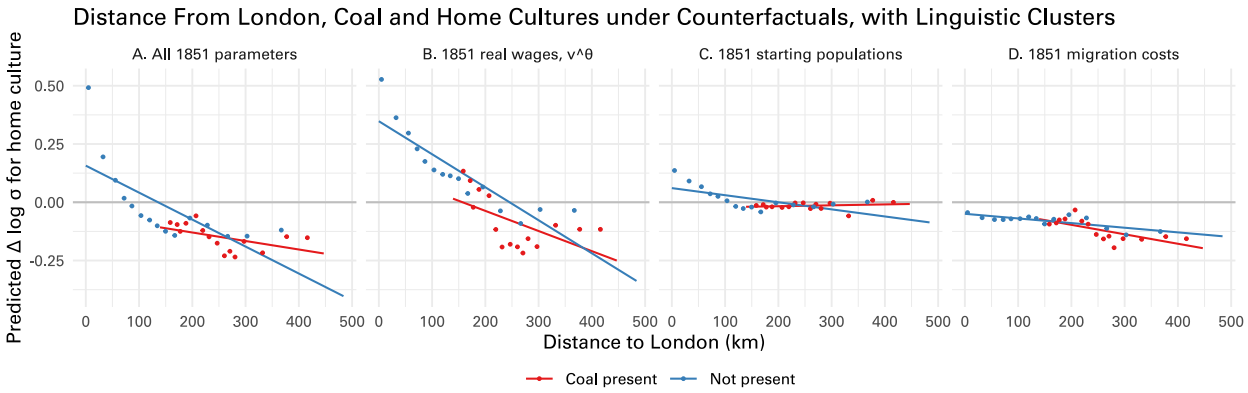


Figure J.11: Average distance, coal, and the changing popularity of home cultures over the late 19th century, with linguistic clusters

This figure replicates Figure 9 using linguistic clusters. The figure shows the predicted change in the log share allocated names most associated with the home culture, across different counterfactual scenarios, subset by whether the district contains a coal deposit. The y axis is the observed log share minus the counterfactual log share, the x axis distance to the City of London. Panel A uses 1851 migration costs, real wages, and starting populations to calculate the counterfactual, B only 1851 real wages, C 1851 starting populations, and D 1851 migration costs.